International Journal of Disaster Risk Reduction 90 (2023) 103651

Contents lists available at ScienceDirect

International Journal of Disaster Risk Reduction

journal homepage: www.elsevier.com/locate/ijdrr

Future exposure modelling for risk-informed decision making in urban planning

Emin Yahya Menteşe [a,*], Gemma Cremen [b], Roberto Gentile [c], Carmine Galasso [b], Maria Evangelina Filippi [d], John McCloskey [e]

[a] Kandilli Observatory and Earthquake Research Institute, Bogazici University, Istanbul, Turkiye
[b] Department of Civil, Environmental and Geomatic Engineering, University College London, London, United Kingdom
[c] Institute for Risk and Disaster Reduction, University College London, London, United Kingdom
[d] School of Sociology, Politics and International Studies, University of Bristol, Bristol, United Kingdom
[e] University of Edinburgh, Edinburgh, United Kingdom

ARTICLE INFO

Dataset link: https://github.com/TomorrowsCities/Tomorrowville

Keywords:
Land use
Plan
Exposure
Disaster risk assessment
Future

ABSTRACT

Population increases and related urban expansion are projected to occur in various parts of the world over the coming decades. These future changes to the urban fabric could fundamentally alter the exposure to natural hazards and the associated vulnerability of people and the built environment with which they interact. Thus, modelling, quantifying, and reducing future urban disaster risk require forward-looking insights that capture the dynamic form of cities. This paper specifically focuses on the exposure component of dynamic natural-hazard disaster risk, by considering urban planning as the centre of future exposure characterisation in a given region. We use the information provided by urban plans and propose an integrated data structure for capturing future exposure to hazards. The proposed data structure provides the necessary detailing for both future physical and socio-demographic exposure in disaster risk modelling. More specifically, it enables users to develop a comprehensive multi-level, multi-scale exposure dataset, characterising attributes of land use, buildings, households and individuals. We showcase the proposed data schema using the virtual urban testbed Tomorrowville. In this case study, we also demonstrate how simplified algorithmic procedures and disaggregation methods can be used to populate the required data. This implementation demonstrates how the proposed exposure data structure can effectively support the development of forward-looking urban visioning scenarios to support decision-making for risk-sensitive and pro-poor urban planning and design in tomorrow's cities.

1. Introduction

In the last sixty years, the world's population has increased from two to eight billion [1] at an unprecedented pace, leading to notable expansion of urban areas. Cities enable people to gain easier access to labour and critical services such as education and health (e.g., [2–4]). However, urban development practices in the last century have often contributed to natural-hazard-induced disaster damages and losses (e.g., [5–8]).

Urbanisation is not slowing down. It is estimated that 70% of the world's population will be living in cities within the next twenty years [1], particularly in low and middle-income countries where the vast majority of urban expansion is expected (e.g., [4,9–12]).

* Corresponding author.
E-mail address: emin.mentese@boun.edu.tr (E.Y. Menteşe).

https://doi.org/10.1016/j.ijdrr.2023.103651
Received 30 September 2022; Received in revised form 6 March 2023; Accepted 19 March 2023
Available online 27 March 2023
2212-4209/

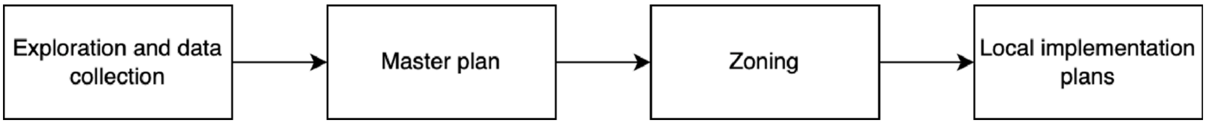

Fig. 1. Phases of urban planning.
Source: Adopted from [39–41].

This will require new structures to be built on undeveloped lands and multiple infrastructure systems to be expanded or improved. These changes will further alter the exposure of people (including their livelihoods and the built environment with which they interact) to disaster risk, and it is critical to understand how [11]. Urban planning offers an opportunity to address this challenge as a forward-looking decision-making process [13] that can contribute to disaster risk prevention. Indeed, urban planning identifies where (and how, to a certain extent) people will construct buildings, locate public spaces, interact with the natural environment and connect assets through networks [14,15].

Within this context, we propose an integrated data structure and potential methods and approaches for exposure modelling that strengthens the link between urban planning and future, people-centred disaster risk-informed decision making facilitated by the Tomorrow's Cities Decision Support Environment (TCDSE) [16,17]. The proposed structure is organised into four layers – (i) land use; (ii) buildings; (iii) households; and (iv) individuals (i.e., people) – that facilitate the creation of detailed future disaster-risk exposure datasets, leveraging basic information provided by urban plans. The configuration of each layer is carefully developed through a comprehensive interdisciplinary process involving relevant experts from urban planning, social science, engineering, and physical science. The proposed data structure ultimately constitutes the spatial (i.e., GIS) backend of the TCDSE Visioning Scenario Module, which produces a detailed characterisation of the future urban system that is subjected to selected multi-hazard simulations as part of future impact and risk assessment procedures in further modules [18].

This paper is structured in four main sections: (i) context and motivation for this study, which involves a brief literature review on urban planning and its interplay with future disaster risk (Section 2); (ii) development of the proposed integrated data structure (Section 3); (iii) demonstration of the structure in disaster-risk assessment, using a virtual urban testbed named Tomorrowville (Section 4); and (iv) conclusions of the study (Section 5).

2. Context and motivation

The urban planning process (Fig. 1) typically starts with an exploration phase, in which relevant stakeholders collect and store pertinent information on the area of interest (e.g., about its population, morphology, geology, the built environment). Then, this information is processed to generate a conceptual master plan that sets the spatial strategies (i.e., where) and policy frames (i.e., how) governing zoning principles. The zoning process subsequently divides the land into several zones representing prominent land uses, such as residential, agricultural, commercial, and industrial (e.g., [19,20]). Criteria such as maximum building height, population density limits, total built-up area, and minimum building plot area are also defined to shape urban development according to a set of principles (e.g., [21,22]). After the zoning phase, specific locations for land uses, such as residential buildings, transportation networks, and public and green spaces, are determined in local implementation plans that provide enhanced detail on how urbanisation will evolve. Across the various phases of urban planning, scenario development is a relatively new approach (mainly adopted in the last decade) that creates an opportunity to engage non-expert participation and fosters visionary thinking (e.g., [23,24]). Scenarios can be regarded as representations of potential urban future alternatives that are conceptualised to test the efficiency of related decisions and policies (e.g., [24,25]). A recent review of scenario development methods can be found in [26]. These methods can be classified into three categories [27]: (1) predictive (forecasting); (2) normative (visioning and backcasting); and (3) exploratory. Predictive scenarios try to estimate the most-likely future configurations of an urban context. Normative scenarios aim for the most desired ones. Exploratory scenarios search for multiple alternative future configurations, accounting for underlying uncertainties [28,29]. There are also hybrid modes of scenario development, where planners rely on flexible methods in highly complex/data-scarce contexts (e.g., [30]). The information provided by these approaches relates to the future form of buildings, infrastructure and socio-economic and demographic characteristics, which are particularly important for assessing the consequences of urban development and enhancing related decision-making processes (e.g., [31,32]). At the core of all scenario-development processes, interdisciplinarity plays an important role that requires interaction, collaboration, and know-how transfer among disciplines [33]. Interdisciplinary approaches in urban planning include both qualitative (e.g., insights, narratives, and social norms) and quantitative (e.g., environmental trends, built environment) aspects (e.g., [34,35]). Despite the effectiveness of scenario development approaches in representing future urban systems, only a few urban scenario studies address future disaster risk (e.g., [36–38]).

It is well recognised that hazards, vulnerabilities and exposure comprising disaster risk change over time (e.g., [42–44]). Yet, their consideration in disaster risk analyses is not future-centric and sufficiently dynamic [45]; current modelling approaches generally estimate risk using a snapshot of the relevant conditions at one point in time. In acknowledged disaster risk assessment frameworks and products such as Hazards United States (HAZUS; [46]), Global Earthquake Model (GEM; [47]), OASIS Loss Modelling

Framework [48], and the Risk Data Library Standard [49], exposure information reflects the present/current situation and has no connection with the future urban context. There are also approaches that try to estimate future urban growth and exposure based on Multiple Agent Systems and Cellular Automata techniques (e.g., [50,51]) for instance, but that disregard urban planning interventions.

This backdrop implies that neither urban planning practice nor disaster risk assessment approaches follow an integrated way of characterising future urban exposure. Hence, the effectiveness of urban planning in reducing disaster risk has not been formally evaluated yet [52–54]. For instance, flood events in Europe in 2013 (e.g., [55,56]) and 2021 (e.g., [57]) and in Australia in 2011 (e.g., [58]) emphasise the existence of gaps in urban planning with respect to disaster risk considerations (e.g., [59,60]). These are partially attributed to the fact that the urban planning process often depends on forecasts driven by past experiences and limited empirical data (e.g., [61]), despite its futuristic core [62]. In addition, expert-oriented and linear (top-down) approaches in planning cannot easily facilitate diverse stakeholders' perspectives and priorities (e.g., [63]). Yet, these are critical for a holistic understanding of disaster risk and its successful integration within the urban planning process (e.g., [64]).

3. Development of an integrated data structure for future exposure modelling

To provide a basis for risk-informed decision making in future urban planning and design, we propose a detailed exposure data structure developed considering the results of a comprehensive interdisciplinary effort [65]. In summary, the knowledge and perspectives of different disciplines helped us to holistically identify and explore the characteristics of an urban system that are relevant for bridging the gap between urban planning and future disaster risk understanding and modelling. The structure represents a future urban scenario by formally capturing four components (i.e., dimensions or layers) of an urban context: (i) land use; (ii) building; (iii) household; (iv) individuals, necessary for people-centred disaster risk quantification.

The proposed data structure leverages a GIS-based approach, where each data layer is kept in a GIS-compatible format either in vector or tabular form. A vector represents a feature using a geometrical object (i.e., point, line, polygon) constructed by one or more connected vertices [66]. A vertex describes a position using coordinates such as X, Y and Z [67]. A single vertex represents a point (e.g., the centre of a building footprint). If the geometry consists of two or more vertices and the first and last vertex are not equal, a polyline feature is formed (e.g., road, river). Where three or more vertices are present, and the last vertex is equal to the first, an enclosed polygon feature (e.g., building footprint) is formed. The vector data form includes an integrated table structure that defines the attributes (fields) of each feature (e.g., a table of building footprint polygons including attributes/fields such as the "number of storeys" and "construction type"). As geometrical attributes (area of a polygon, coordinates of the vertices and points) are stored automatically at the backend in GIS, they are excluded from the final proposed data structure. The household and individual information are stored in tabular format on a relational basis.

The exposure dataset is produced for a time in the future t_f according to the proposed data structure, through the interdisciplinary process depicted in Fig. 2. First, the planning extent is selected and subdivided into zones that define the aggregated projected land-use types (stored in the land-use plan layer). The land-use information is used in conjunction with assumptions on future hazards, buildings/physical infrastructure, and socio-economic characteristics to generate spatial information on building locations and their attributes (stored in the building layer) as well as on households living in those buildings. The building layer is enriched with building taxonomy information to facilitate the physical infrastructure impact quantification within disaster risk assessments. Socio-economic and demographic projections are used for characterising data on households (stored in the household layer) and individual people within each household (stored in the individual layer).

Exposure development based on the proposed data structure is a hybrid scenario generation process, combining the predictive and exploratory scenario development approaches described in Section 2. It is predictive as we attempt to project physical and socio-economic/demographic attributes in a realistic way. It is exploratory as we create different future exposure scenarios to cover a wide range of urban development options that may occur. This hybrid scenario development approach is useful in disaster risk prevention, in which decision-making requires an awareness of all of the uncertainties involved. There is no expectation that any particular projection will be realised, but the uncertain scenarios produced are useful for relative comparisons of the associated disaster risk consequences. In the following sub-sections, we provide details on what the proposed data structure must include by means of attributes (Sections 3.1 to 3.4) and how these attributes can be generated based on several techniques (Section 3.5).

3.1. Land-use plan layer

Land-use zoning can be regarded as the main task of urban planning, in which stakeholders (e.g., governments, municipalities, communities, researchers, private sector) decide the future functionality of a given space for a certain period (e.g., [68]). A land-use plan layer typically includes attributes on the type of use and the associated maximum population densities, floor area ratio and/or setback distances (e.g., [69]) of the given space. These attributes determine how and by how many people an urban space can be inhabited in the future [70]. The land-use plan layer of our proposed data structure broadly follows this convention, incorporating some supplementary attributes related to socio-economic status (SES) and demographic characteristics. SES-related attributes can be defined based on the acknowledged literature that suggests income, education and occupation as the main components defining SES (e.g., [71]). Among these three attributes, we use income level as a proxy to represent SES within the boundaries of a land-use type, because income-based segregation within residential areas is quite common (e.g., [72–75]). The proposed land-use attributes and their definitions are given in Table 1.

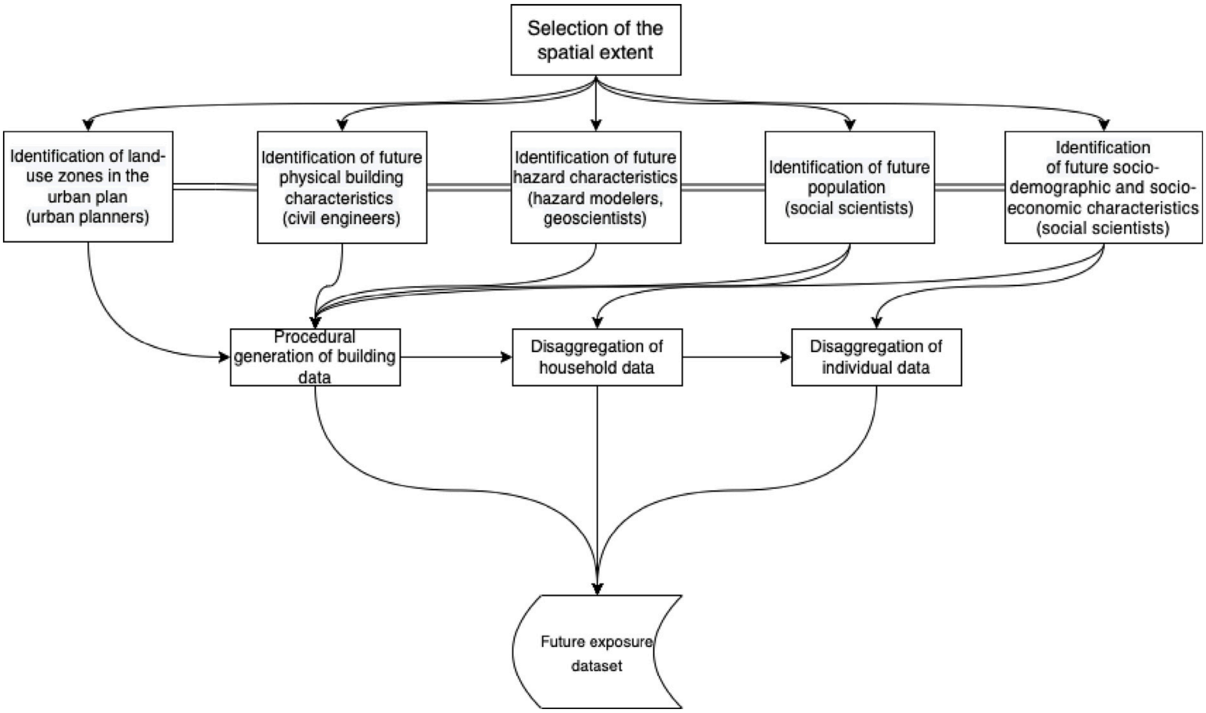

Fig. 2. Development of exposure data, based on the proposed data structure.

Table 1
The attribute table of the land-use plan layer. Unless specified otherwise, all attributes correspond to time t_f.

Field	Alias	Definition	Type
zoneID	Zone ID	Unique land-use zone (polygon) identifier number	Integer
LuF	Proposed land-use type	Land-use type of the zone	String
population	Initial population	Number of people inhabiting the zone at the time of generating the land-use plan (t_0)	Integer
densityCap	Maximum density capacity	Maximum number of persons per hectare of the zone (density capacity of the zone)	Float
setback	Setback distances	Minimum distance of the building footprint boundaries from the outer land parcel boundaries or a road	Float
floorAreaRatio	Floor area ratio	Ratio of building plot over building's parcel area (land plot)	Float
avgIncome	Average income level	Average income level of households residing within the zone	String

3.2. Building layer

The proposed building layer includes a list of the minimum attributes required for quantifying related future natural-hazard-induced impact metrics. Each building is assigned a unique building ID linked to the land-use layer through the "zoneID" field, which contains the ID of the land-use zone (polygon) hosting the building. The attribute "specialFac" identifies the special facility status of the building; it is an integer equal to zero for "standard" occupancy types (i.e., residential, commercial, industrial) and greater than zero for special occupancy types (e.g., 1 for schools, 2 for hospitals). The "repValue" attribute indicates the absolute economic cost of replacing the building in case of complete reconstruction. This attribute could incorporate t_f-dependent price fluctuations (e.g., inflation) if necessary for the application of interest, which would need to be applied in a consistent manner across all buildings to enable relative comparisons of building values. For buildings with residential occupations, the "nHouse" and "residential" attributes respectively indicate the number of households and residents within the building. These two attributes are equal to zero for non-residential buildings; the theoretical maximum number of people that would occupy these buildings at any point can be approximated as the total number of individuals that rely on them (identified through the "communityFacID" of the household layer and the "indivFacID" of the individual layer — see Sections 3.4 and 3.5). (Note that temporal variations in building occupancy are not explicitly accounted for within the current version of the proposed data structure, but will be integrated as part of a future update.)

Physical attributes of each building related to the physical impacts of natural hazards are condensed in the exposure taxonomy string attribute (Field:expStr). The global exposure database for all (GED4ALL) [76] taxonomy string format is used for the proposed data structure, since it facilitates the consideration of different asset classes (i.e., buildings, roads, railways, bridges, pipelines,

Table 2
Complete set of GED4ALL attributes for buildings exposed to different hazards. OSM: OpenSteetMap. (Modified after [78]).

Attribute	OSM Key
Direction	building:direction
Material of LLRS	building:lateral:material
Lateral Load Resisting System (LLRS)	building:lateral:system
Height	building:levels
Date of construction or retrofit	building:age
Surroundings	building:adjacency
Occupancy	building
Shape of building plan	building:shape
Structural irregularity	building:irregularity
Ground floor hydrodynamics	ground_floor
Exterior walls	building:material
Roof shape	roof:shape
Floor system material	floor:material
Foundation	building:foundation
Fire protection	building:fireproof

Table 3
The attribute table of the building layer. All attributes correspond to time t_f.

Field	Alias	Definition	Type
zoneID	Parent land-use zone ID	Unique ID number of the land-use zone where the building is located	Integer
bldID	Building ID	Unique ID number for each building	Integer
specialFac	Special facility status	Number corresponding to the "special facility" status of the building	Integer
repValue	Replacement value	Total replacement value of the building	Float
nHouse	Number of households	Number of households residing in the building	Integer
residents	Residents	Number of individuals residing in the building	Integer
expStr	Exposure taxonomy string	Taxonomy string based on GED4ALL	String

storage tanks, power grids, energy generation facilities, crops, livestock, forestry, and socio-economic data), and different hazards (i.e., earthquakes, volcanoes, floods, tsunamis, storms, cyclones and drought). Table 2 provides the GED4ALL attributes for buildings, also highlighting how they are mapped to the keys (or tags) in the OpenStreetMap database [77], a large open-source mapping repository.

For example, the lateral load resisting system attribute is used to define physical vulnerability to earthquakes and wind (among other hazards), the presence of a basement affects flooding vulnerability, and the roof typology (and its features) influences windstorm vulnerability. Each attribute of the GED4ALL string can be defined with different levels of refinement (i.e., from Level 1 to Level 3) to accommodate various levels of data availability. An example of a Level 1 string is CR/H:1/LWAL/IND, which indicates a one-storey reinforced concrete industrial building with a wall lateral load-resisting system. Level 2 information could include data on the material technology, for instance, which may be cast in place (CR+CIP) in the above example. A Level 3 string could include secondary information related to primary (Level 2) structural irregularities, such as the presence of torsion eccentricity and a re-entrant corner for the above example (IRIR+IRPP:TOR+IRPS:REC). For each asset class, detailed documentation is provided at [79]. The GED4ALL string easily links the specific combination of building attributes to a specific physical impact model (single- or multi-hazard-fragility or vulnerability relationships). Specific details on how to characterise this link are provided elsewhere [78]. The attribute table of the building layer is provided in Table 3.

3.3. Household layer

This layer represents the social connections of individuals who are members of the same households and can be leveraged to capture their collective experience of a natural-hazard-related disaster. This layer is interconnected with the underlying residential building information through the "bldID" attribute. This field also makes it possible to determine which land-use zone the household is living in and what type of socio-economic status is dominant in the area (Table 4).

Indicating a household's income level (Field:income) enables impact metrics to be disaggregated based on the socio-economic characteristics of affected people. This is essential for highlighting disproportionate impacts of natural-hazard events that are often experienced by the poorest households. Low-income households are generally exposed to and affected by hazards more than wealthier ones, since they lose a more significant portion of their income and assets in case of a disaster (e.g., [80,81]) and have fewer resources to recover after a disaster event. Note that the "income" attribute is a relative categorical variable, and therefore implicitly accounts for t_f dependent inflation, for instance. The layer includes information on the closest critical facilities to the household (Field:communityFacID) – which could include medical facilities, schools, or grocery stores, for instance – to help determine consequences that arise from a lack of accessibility to these services after a disaster [17].

Table 4
The attribute table of the household layer. All attributes correspond to time t_f.

Field	Alias	Definition	Type
bldID	Parent building ID	ID Unique ID number of the residential building in which the household lives	Integer
hhID	Household ID	Unique ID number for each household	Integer
income	Household income	String indicating the level of income of the household	String
nInd	Number of individuals	Number of individuals living in the household	Integer
communityFacID	Community facility ID	Unique ID numbers of the closest critical facilities (e.g., hospital, grocery store, fire station) to the household's residence	Vector

Table 5
The attribute table of the individual layer. All attributes correspond to time t_f.

Field	Alias	Definition	Type
hhID	Parent household ID	Unique ID number of the household to which the individual belongs	Integer
indivID	Person ID	Unique ID number of the individual	Integer
gender	Gender	Number indicating the gender of the individual	Integer
age	Age	Number indicating the age range of the individual	Integer
head	Head of household status	Number indicating whether this individual is the head of household	Integer
eduAttStat	Education attainment status	Number indicating the education attainment level of an adult individual	Integer
indivFacID	Individual facility ID	Unique ID numbers of the buildings that the individual regularly visits (can be workplace, school, etc.)	Vector

3.4. *Individual layer*

Documenting the characteristics of future individual members of an urban system is critical for quantifying people-centred disaster-risk impacts, such as unemployment, displacement, and inaccessibility to education [17]. The attribute table of the individual layer (Table 5) includes information on gender (Field:gender), age (Field:age), educational attainment (Field:eduAttStat) and head-of-household status (Field:head), which are widely used indicators of social vulnerability (e.g. [82–85] that can be leveraged to distinguish a person's general reliance on the built environment. For instance, a child may depend on a school, a working adult of high educational attainment may work in an office, and a woman of low educational attainment who is not the head of the household may spend a significant amount of time at their residence. These dependencies – which are documented in terms of individual buildings within the attribute table (Field:indivFacID) – can then be used to characterise the consequent disruption to daily lives that would result from disaster-induced damage to the built environment. Note that these individual facilities are different from the buildings provided for "communityFacID" in the household layer; meaning that these buildings are specifically relevant for a given individual and do not necessarily serve the same household members.

3.5. *Potential data generation approaches*

It is important to acknowledge that the detailed projections on discrete buildings, households, and individuals required to populate the proposed data structure are unlikely to be available for most urban contexts. However, disaggregation of available coarser projected data into separate buildings, households, and individual units can be achieved using various techniques.

Procedural modelling (PM), which is defined as the development of content through a procedure or a set of prescribed parameters [86,87], enhances the spatial details of urban plans based on criteria defined in master planning and zoning processes (e.g., [88]). PM is increasingly being adopted in urban planning practice (e.g., [89–92]). The creation of urban layouts via procedural algorithms was first introduced by [93], which leverages geography, population density and geographical extents as input data. Agent-based modelling approaches are also used in PM to generate building layout patterns (e.g., [94]) and to guide interactions that ultimately lead to the creation of land use zones such as residential, commercial, industrial, recreation, and roads (e.g., [95]). PM can also be used to generate terrain, vegetation and water bodies that are essential for comprehensive characterisation of a spatial context [96]. A detailed, relatively recent literature review on PM can be found in [97].

Although not explicitly connected with urban planning practice, disaggregation of coarse exposure datasets based on ancillary information, also called "dasymetric mapping" (e.g., [98]), is a common approach used to enhance the resolution of urban-layout data within quantitative disaster risk assessments [37]. It involves transforming aggregated information (i.e., the population at the city level) into more local units (i.e., neighbourhood level) that capture the spatial distribution of information in finer detail [99]. In a recent related study across Europe by [100], it is found that disaggregating exposure information to finer resolution increases the accuracy of earthquake loss estimation analysis.

Synthetic population generation algorithms, which are widely employed as part of different applications involving space-dependent data for individual decision-making (e.g., mobility, disease spreading, population energy demand, and population

Table 6
Assigned attribute values in the land-use plan layer.

LuF	Population	densityCap	avgIncome
Agriculture	1308	40	lowIncomeA
CityCentre	1049	300	midIncome
CommercialResidential	556	250	midIncome
CommercialResidential	1107	250	midIncome
CommercialResidential	443	250	lowIncomeA
CommercialResidential	273	250	lowIncomeA
HistoricalPreservation	914	350	na
Industry	0	125	na
Recreation	0	0	na
Residential (Gated Neighbourhood)	1297	100	highIncome
Residential (High-Density)	5437	350	lowIncomeB
Residential (High-Density)	4069	350	lowIncomeA
Residential (Moderate-Density)	2028	200	highIncome
Residential (Moderate-Density)	11 776	250	midIncome
Residential (Moderate-Density)	6309	250	lowIncomeB
Residential (Low-Density)	2501	100	highIncome

growth), can be used to translate micro samples of buildings, households, or individuals and account for marginal distributions of their attributes in relevant synthetic datasets. Furthermore, activity-based (or human mobility) models, which represent a person's daily travel demand, can be leveraged to capture the dynamic dependencies of individuals on the built environment and determine their interaction with physical assets. Interested readers are referred to Appendix A (Table A.8) and Appendix B (Table B.9) for more details on synthetic population generation and activity-based modelling, respectively.

4. Case-study demonstration — Tomorrowville

The case-study demonstration in this section showcases how the proposed future exposure data structure can be adopted for a Global South data-scarce context, where simplistic ad-hoc procedures are necessary for generating the required data. We use a 500 ha virtual urban testbed called Tomorrowville, which represents a Global South urban setting based on Nairobi (Kenya) and Kathmandu (Nepal) contexts, and is currently (at t_0) home to 39,058 people. The exposure dataset is generated for a time 50 years in the future (i.e., $t_f = t_{50}$) when Tomorrowville is expected to cater for a population of approximately 80,000 people.

Producing Tomorrowville relied on a seven-month-long interdisciplinary scenario development process involving 19 researchers and experts from different disciplines, such as engineering, physical science, risk modelling, social sciences and urban planning. This process ultimately defined assumptions on future urbanisation patterns [65], which were used to populate the urban characteristics of Tomorrowville visioning scenarios based on the proposed exposure data structure. These assumptions are based on relevant studies of Kathmandu and Nairobi [101–106] as well as expert judgement. Note that some attributes within the proposed data structure are neglected for this application, given the underlying data scarcity (highlighting the ability of the proposed structure to adapt to specifically challenging circumstances).

4.1. Land-use plan layer

Future land-use types within Tomorrowville address pertinent demands for housing, workplaces, and public services (e.g., school, hospital), and facilitate agricultural areas as well as natural environment spaces (i.e. forests). There is also a historical preservation area with a unique socio-cultural significance, representing an "old-town" settlement. Assignment of attribute values for the land-use plan layer are summarised in Table 6. (Note that the "floorAreaRatio" and "setback" fields are not used in the case study, since Tomorrowville does not have defined land plots in yet-to-be developed areas).

Residential zones are classified on the basis of population density into four land-use types: (i) "low-density residential"; (ii) "moderate-density residential"; (iii) "high-density residential"; and (iv) "gated neighbourhood". The "avgIncome" field has four possible values: (i) "highIncome"; (ii) "moderateIncome"; (iii) "lowIncomeA"; and (iv) "lowIncomeB", where "lowIncomeA" represents a relatively higher income level than "lowIncomeB". The "CityCentre" zone is located close to the high-income residential areas. It includes commercial and governmental uses as suggested by urban planning and social scientists to represent a business district within Tomorrowville. The "CommercialResidential" land-use type represents a mixed occupation, including both commercial and residential uses. It reflects the expert recommendations of local Kathmandu researchers in the interdisciplinary team, who noted that buildings consisting of commercial units on the first two floors and residences on higher floors are particularly common in the Kathmandu region. Each zone and its avgIncome value (where relevant) are shown in Fig. 3. Tomorrowville datasets can be found at https://github.com/TomorrowsCities/Tomorrowville.

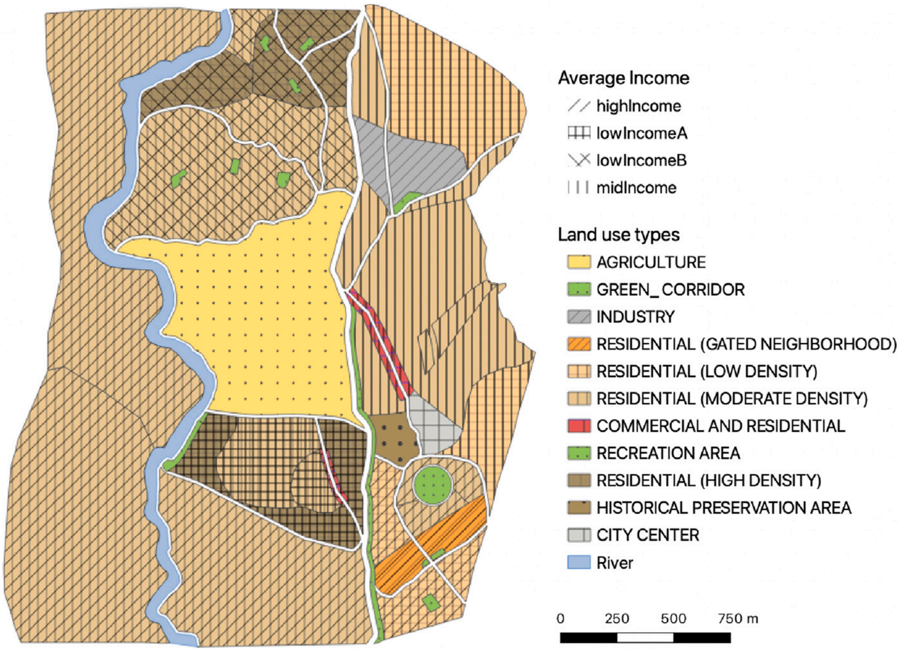

Fig. 3. Tomorrowville land-use plan and projected income levels.

Table 7
Input variables for the residential building data generation algorithm.

Variable explanation	Symbol	Applies to	Source
Current population	$P_{current}$	Whole area	User assumption
Projected population at t_f (i.e., 50 years in the future in this case)	P_{target}	Whole area	User assumption, consistent with “densityCap” field values
Current population density	$d(poly)$	Each polygon	User assumption
Population density capacity	$d_{cap}(poly)$	Each polygon	“densityCap” field values
Area	$A(poly)$	Each polygon	Geometrical attributes of the land-use plan layer
Building footprint area probability distribution	$A_b(poly)$	Each polygon	User assumption
Pertinent characteristics (e.g., average, variance, …) of $A_b(poly)$	$\bar{A}_b(poly)$; $\sigma^2(A_b)(poly)$	Each polygon	User assumption
Average number of storeys	$N_s(poly)$	Each polygon	User assumption
Average building area per household	$\bar{A}_h(poly)$	Each polygon	User assumption
Average household size	$\bar{N}_{hc}(poly)$	Each polygon	User assumption

4.2. Building layer

Three building layers have been created for Tomorrowville. The first building layout ($TV0_{b0}$; Fig. 4) refers to the present-day configuration of Tomorrowville (at t_0), which includes 4,810 buildings. Approximately 60% of these buildings belong to “high density” land-use zones (polygons), 13% are in “medium density” ones, 14% belong to “low density” ones, and the remaining 13% are spread across other land-use zones. To reflect a Global South context, $TV0_{b0}$ buildings primarily consist of low-rise (1–4 storey) masonry buildings. Reinforced concrete structures constitute 27% of buildings, 85% of which are mid-rise (4 to 9 storeys) (e.g., [104,105]). There are also adobe and stone buildings located in low-income zones.

The other two building layers ($TV50_{b1}$ and $TV50_{b2}$) represent different possible configurations of the buildings to be built within the next 50 years (described below and shown in Fig. 5). To create these building layouts, we developed a data generation algorithm to: (i) synthetically generate a set of buildings consistent with an assumed population demand; and (ii) assign building-by-building attributes (see Tables 2 and 3) based on coarser data (e.g., land-use data) and a set of assumed input parameters (Appendix C, Table C.10). The algorithm requires various inputs that may be related to the entire urban area, the “avgIncome” value of the associated land-use zone (polygon), or the “LuF” value of the associated land-use zone. The input variables, and their spatial scope of application, are detailed in Table 7.

Fig. 4. $TV0_{b0}$ building layout.

The first step of the algorithm involves calculating the "effective" population density $d_{cap}^{eff}(poly)$ of each land-use zone (polygon) according to Eq. (1).

$$d_{cap}^{eff}(poly) = d_{cap}(poly) - d(poly) \tag{1}$$

Given the current population of the case-study area ($P_{current}$) based on the "Population" field values, the additional population to allocate in the considered scenario is equal to $P_{target} - P_{current}$, where P_{target} = 80,000. Each polygon is allocated a share of population, $p_{alloc}(poly)$, proportional to $d_{cap}^{eff}(poly)$:

$$p_{alloc}(poly) = (P_{target} - P_{current})\frac{d_{cap}^{eff}(poly)A(poly)}{\sum_{poly} d_{cap}^{eff}(poly)} \tag{2}$$

The algorithm involves using Monte Carlo sampling to generate, for each polygon, residential buildings of a random footprint area $A_b(bld)$ according to a selected distribution ($p(A_b(poly))$) until $p_{alloc}(poly)$ is met. This condition is monitored by calculating the number of households allocated in each generated residential building $h_{alloc}(bld)$ (Eq. (3)) and keeping track of the cumulative sum of allocated people, $p_{alloc}(bld)$ from the average household size $\bar{N_{hc}}(poly)$.

$$h_{alloc}(bld) = \text{round}\left(\frac{A_b(bld)N_s(poly)}{\bar{A_h}(poly)}\right) \tag{3}$$

Based on expert judgement within the interdisciplinary group, $p(A_b(poly))$ are assumed to be uniform distributions for all residential buildings in this specific case study: the limits of the distributions for residential buildings are equal to [32,66] m^2 for land-use zones with "lowIncomeA" and "lowIncomeB" avgIncome field values, [32,78] m^2 for zones with "midIncome" avgIncome field values, and [70,132] m^2 for zones with "highIncome" avgIncome field values. $\bar{A_h}(poly)$ values are 44 m^2, 54 m^2, and 67 m^2, respectively,

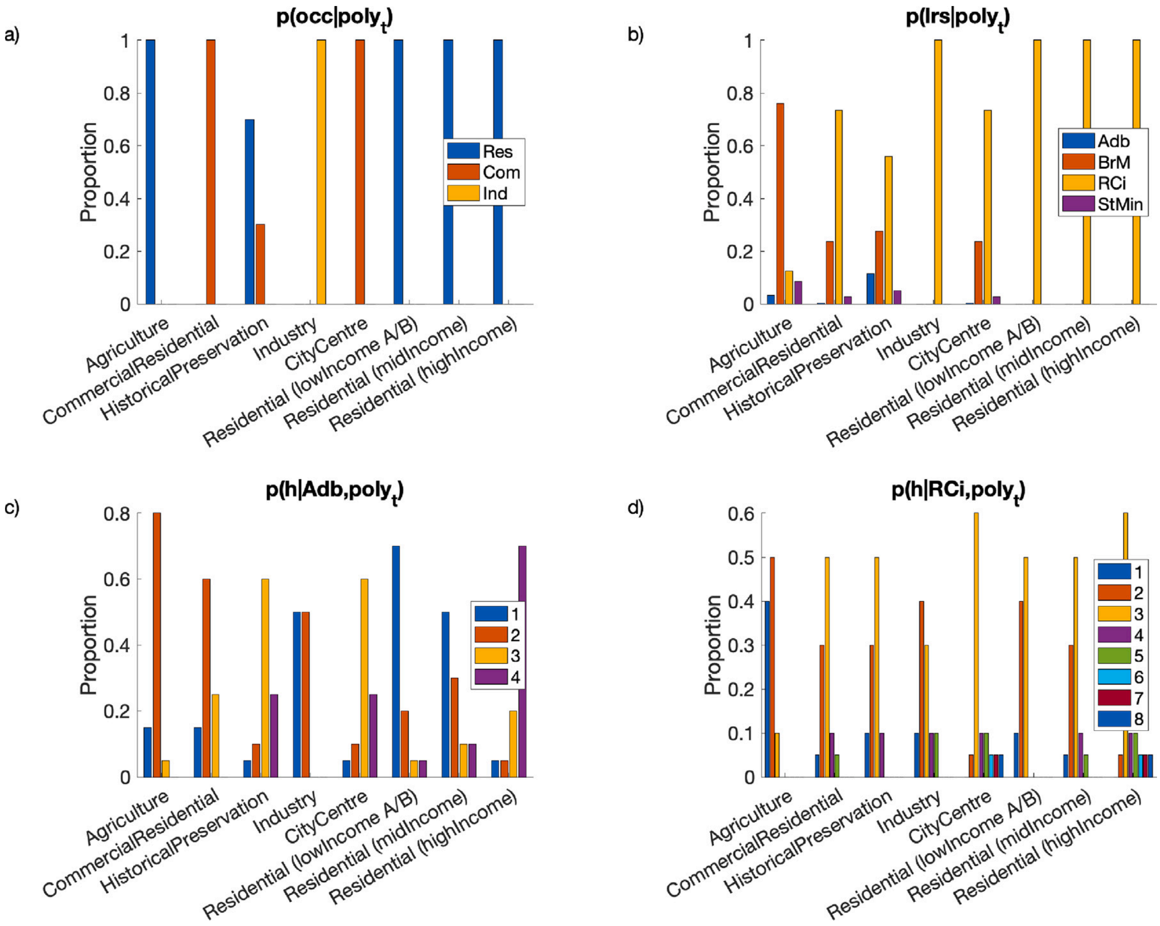

Fig. 5. Adopted distributions for the building data generation algorithm: (a) $p(occ|poly_t)$; (b) $p(lrs|poly_t)$; (c) $p(h|Adb, poly_t)$, also valid for BrCfl, BrCri, BrM; (d) $p(h|Rci, poly_t)$. Res: residential; Com: commercial; Ind: industrial; BrCfl: brick and cement with flexible floor; BrCri: brick and cement with rigid floor; BrM: brick and mud.

for polygons with "lowIncomeA/B", "midIncome", and "highIncome" avgIncome field values. $\bar{N_{hc}}(poly)$ is respectively equal to 2, 4, and 3 for polygons with "lowIncomeA/B", "midincome", and "highIncome" avgIncome field values.

To complete the definition of the $TV50_{b1}$ and $TV50_{b2}$ residential building layers according to Section 3.2 (Fig. 6), each generated residential building must be assigned a GED4ALL taxonomy string so that it is possible to associate it with relevant multi-hazard physical impact models [78]. The next steps of this algorithm use Monte Carlo sampling to randomly characterise the generated residential buildings accordingly. The variables involved in this sampling are the various parameters of the GED4ALL taxonomy strings. Based on the available data, we describe the taxonomy strings using four main attributes: occupancy type (occ), construction material and lateral load resisting system (lrs), height (h), and design code level ($code$). An extension of the presented procedure to include more attributes is straightforward.

The next steps of the algorithm are as follows:

- For each unique combination of general land-use type and "avgIncome" that features across the prescribed land-use zones (see Table 6) – denoted as $poly_t$ – assign the distribution of occupancy type, $p(occ|poly_t)$ (Fig. 5a);
- For each $poly_t$, define the distribution of the building construction material and lateral resisting system, $p(lrs|poly_t)$. The lrs distributions adopted for both $TV50_{b1}$ and $TV50_{b2}$ are shown in Fig. 5b;
- For each unique combination of lrs and $poly_t$, define the distribution of the building heights, $p(h|lrs, poly_t)$. The adopted distributions for both $TV50_{b1}$ and $TV50_{b2}$ are shown in Fig. 5c for the Adb, BrCfl, BrCri, BrM buildings and Fig. 5d for the RCi buildings;
- For each unique combination of lrs and $poly_t$, define the distribution of the building design code level (low, moderate, high), $p(code|lrs, poly_t)$. For Tomorrowville, this is achieved by mapping data on year of construction to the evolution of the building design code in Kathmandu (as detailed in [78]);
- For each polygon and its generated residential buildings: (i) sample the parameters $lrs, h, code$ using the above distributions; and (ii) define and assign the corresponding GED4ALL string (see [78]).

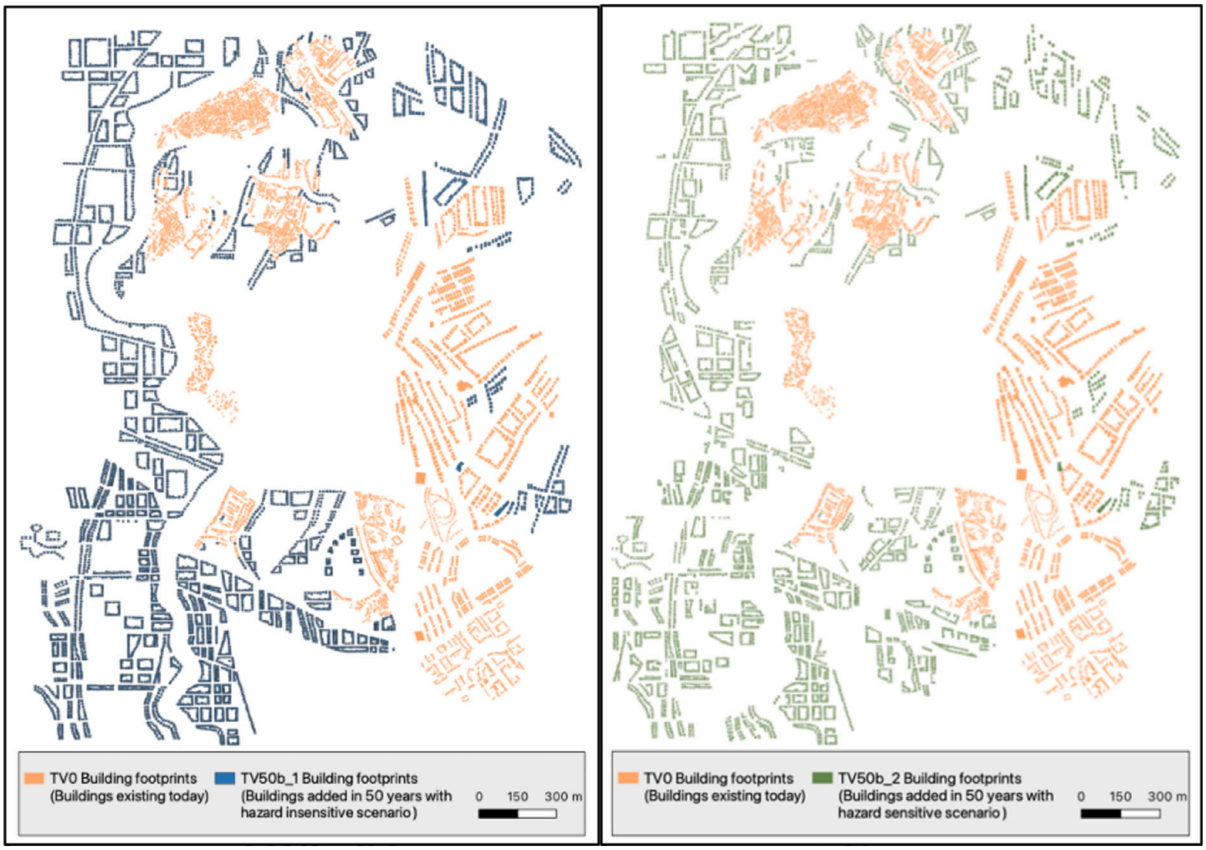

Fig. 6. TV50 buildings added to $TV0_{b0}$ (resulting in a total of 10,156 buildings) in the (a) $TV50_{b1}$ scenario, and the (b) $TV50_{b1}$ scenario.

Non-residential (i.e., commercial and industrial) buildings are generated until they comply with the $p(occ|poly_t)$ distributions and assumptions for TV50 specified in Table C.10 and shown in Fig. 5a. $p(A_b(poly))$ uniform distribution limits for commercial and industrial buildings are [12,88] m^2. These buildings are assigned GED4ALL strings according to the procedure described for residential buildings.

The next step of the algorithm assigns a location to the generated buildings within each polygon. This may be determined from the urban design process itself (conditioned on criteria related to distances between buildings, their potential connection through transportation and the digital elevation model of the area), and is therefore not prescribed in detail here. For the specific case of $TV50_{b1}$, we first randomly allocated buildings within their corresponding polygon, and finalised an ad-hoc urban layout by manually shifting and rotating the building footprints in the GIS environment, in line with the advice of urban planning experts on the interdisciplinary team. We also manually added two hospitals and ten schools. $TV50_{b2}$ represents a rearrangement of the buildings within $TV50_{b1}$ to avoid development in areas affected by flooding and debris flows [107] (see Fig. 6).

4.3. Household layer

Information sources used to generate the household (and individual) data comprise a combination of expert judgement and geographically relevant literature and are described in Appendix D.11 (which also provides the values for each variable). Given the lack of detailed information, we use a straightforward Monte Carlo sampling strategy to sample uncertain household and individual attributes. It is assumed that the population across Tomorrowville (and each land-use zone that includes residential buildings) is evenly split between male and female genders, in the absence of more detailed data. Field:nInd is first sampled according to a distribution conditional on the avgIncome field value of the land-use zone in which the household is located, $p(nInd|avgIncome)$ (the "income" attribute' is neglected in this case study, since household incomes are assumed not to vary significantly with respect to the corresponding avgIncome assignment). The head of each household is assigned a gender ($gender_{hh}$) according to the conditional distribution, $p(gender_{hh}|avgIncome)$. The age of the head of household (age_{hh}) is sampled according to the conditional distribution, $p(age_{hh}|avgIncome, gender_{hh})$, where age_{hh} accounts for only those old enough to be head of household ($age_{hh,min}$). Each remaining person in a household is equally likely to be of either gender and has an age that is sampled according to the conditional distribution, $p(age|avgIncome, gender)$. Thus, there are no explicit dependencies between the age and/or gender of members of the same household. Field:communityFacID is assumed to comprise the nearest hospital building.

4.4. Individual layer

The age and gender field values are assigned as discussed in Section 4.3. The "eduAttStat" attribute (possible values are "none": no schooling, "primary": primary schooling only, "secondary": secondary schooling only, and "university": college- or university-level education) is sampled independently for all adults (i.e., for which $age \geq age_{adult}$; see Appendix D, Table D.11). Adults with $eduAttStat = university$ less than or equal to the oldest age at which a person typically attends university ($age_{university}$) are assumed to attend university or some other tertiary institute outside of Tomorrowville. Field:indivFacID covers workplaces and schools in this case study.

All non-college-attending adults less than a retirement age that is specific to the avgIncome field value of the land-use zone in which they reside (i.e., age_{old}; see Appendix D, Table D.11) are assigned a broad workplace location (wp) category (i.e., "$poly_t$" value, but including the possibility of unemployment) according to distributions conditional on gender, the avgIncome field of the land-use zone in which they reside, and the educational attainment rate of the individual adult $p(wp|gender, avgIncome, eduAttStat, age_{adult} \leq age \leq age_{old})$. The individual is equally likely to work within any of the workplace buildings associated with the assigned workplace location category.

Children of school-going age (i.e., individuals with $age_{school_{min}} \leq age \leq age_{school_{max}}$, where $age_{school_{min}}$ and $age_{school_{max}}$ vary depending on the avgIncome field of the land-use zone in which they reside) are assumed to attend the nearest school according to gender-specific statistics that also change according to the avgIncome field of the land-use zone in which they reside as well as the eduAttStat field value of their head of household $eduAttStat_{hh}$, $p(school|age_{school_{min}} \leq age \leq age_{school_{max}}, gender, avgIncome, eduAttStat_{hh})$; see Appendix D, Table D.11.

5. Conclusions

This paper introduced a comprehensive future exposure data structure that can be used to strengthen the link between traditional urban planning practice and forward-looking, people-centred disaster risk decision support studies [18]. The proposed structure facilitates spatial links between physical and social exposure through a multi-layered GIS database that accounts for varying levels of detail on the future urban system. It can also be leveraged for a holistic exploration of the consequences of risk-related policies on future urban development. Suggested procedures for generating the high-resolution data to be stored in the proposed structure were also provided. The proposed approach represents a significant contribution to (1) identifying, understanding and using current and future disaster risk scenarios; and (2) facilitating the pursuit of resilient urban development and design, which are two of the United Nations Office for Disaster Risk Reduction Ten Essentials for Making Cities Resilient [14,108] a campaign set up to accelerate the implementation of the Sendai Framework for Disaster Risk Reduction [5] at local level.

We demonstrated an implementation of the proposed data structure for a virtual urban testbed known as Tomrrowville, which has been designed to represent a generic Global South city. The data-scarce circumstances of the case study required the development of ad-hoc algorithms to procedurally generate the required physical and social data within the specified urban extent. The exposure database developed for Tomorrowville in this paper is leveraged for risk-informed decision support on future urban development in [18], demonstrating the significant utility of the proposed structure in a forward-looking natural-hazard risk assessment context.

While the proposed data structure fills a notable gap in the state-of-the-art regarding future risk exposure characterisation, there are a number of limitations associated with it. Firstly, the data structure only accounts for separate (nodal) physical and social entities (e.g., buildings and people); future work will establish additional layers of the GIS database that characterise the important networks that connect (and interconnect) these nodes. Capturing these networks will enable a more enriched quantification of future natural-hazard impacts, including the inaccessibility experienced by people to significant physical (i.e., schools, hospitals) and social (i.e., friends' and relatives' houses) focal points within the urban system. Furthermore, our research will investigate incorporating the dynamic nature of certain attributes (e.g., population, income level, replacement value) in a more structured way within the exposure modelling process. A comprehensive visualisation platform for the stored exposure will also be created, transforming the proposed data structure into a powerful tool that can guide and inform relevant stakeholders on risk-informed future urban development.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The exposure dataset for the Tomorrowville virtual urban testbed can be found at: https://github.com/TomorrowsCities/Tomorrowville

Acknowledgements

The authors acknowledge funding from UKRI GCRF under grant NE/S009000/1, Tomorrow's Cities Hub and the invaluable comments made by the anonymous reviewers who helped us to improve our paper.

Appendix A. Literature review of synthetic population generation approaches

Table A.8 provides a summary of popular synthetic generation methods that have been employed in the literature, along with their advantages and disadvantages. While these methods have been traditionally leveraged for generating synthetic populations of households and individuals (and are described as such in the following tables), non-hierarchical models for generating individuals (e.g., iterative proportional fitting, Monte Carlo Simulation methods, combinatorial optimisation) could also be used to generate buildings if the necessary information (e.g., marginal distributions of building attributes and relevant inter-attribute correlation data) is available.

Appendix B. Literature review of activity-based modelling approaches

Activity-based (or human mobility) models seek to represent the travel demand of an individual over an entire 24-h period, including the sequence of activities and (in some cases) the exact locations and durations of these activities. As such, they are useful for characterising the dependencies of people on the built environment (i.e., for populating the FacilityID attributes of the individual layer). Activity-based models may be broadly categorised as [122]: (1) constraint-based (which assign trips based on their feasibility within a space–time context); (2) utility-maximising (which focus on the activity-related choice-making of agents using a pre-specified utility function); and (3) computational process models (rule-based approaches that capture decision heuristics of individuals, for more natural decision making under uncertainty than that of the models in category (2)). Depending on their categorisation, these models can require a significant amount of information, ranging from origin–destination matrices (e.g., [123]) to a travel diary/survey that records the movement and activity participation of people over a given day (e.g., [124–126]), urban sensing (e.g., mobile phone) data that contains intrinsic mobility patterns (e.g., [126–130]). Recent efforts have focused on exclusively leveraging open-access data [131] to fulfil the necessary information requirements. Given the context of their traditional application to transport demand modelling, most activity-based models particularly focus on the structure of daily patterns (i.e., the sequences or patterns of activities and resulting trip chains) rather than the description of exact locations that individuals frequent within the urban system (e.g., [128,132]). We focus here on the relatively few efforts of the literature that have centred on the latter challenge, which is particularly relevant for the scope of this work. A brief literature review of a small selection of these studies is provided in Table B.9; the studies were specifically chosen to capture a variety of activity-based model categories, considered locations, (mathematical) methods used, and required information.

Appendix C. General assumptions for building data generation in Tomorrowville

Kathmandu census data were used to determine six material types (adobe, brick in cement with rigid floor, brick in cement with flexible floor, brick in mud, reinforced concrete, stone in mud) and their related distributions for each $poly_t$. A combination of census data and qualitative expert knowledge of local researchers were then used to define eight unique values of the building taxonomy strings. The $p(lrs|poly_t)$ distributions were constructed by assuming that every adobe or brick building is composed of walls, and every reinforced concrete building is composed of masonry-infilled frames. The $p(code|lrs, poly_t)$ distributions were constructed by: (i) assuming that all wall masonry buildings are low code; and (ii) using the distributions of $p(code|lrs)$ provided by the Kathmandu census data for the reinforced concrete frame buildings. The $p(occ|poly_t)$ distribution is summarised in Table C.10. More detailed information regarding building attributes and the distribution of taxonomy strings is provided by [78]. Further assumptions are:

1. All buildings within $TV0_{b0}$ still exist after 50 years.
2. More than 95% of buildings constructed within $TV0_{b1}$ or $TV0_{b2}$ are reinforced concrete with infills.
3. $TV0_{b1}$ and $TV0_{b2}$ do not feature any stone or adobe buildings.
4. Buildings in $TV0_{b1}$ and $TV0_{b2}$ are either low-rise (1–4 storeys) or mid-rise (5–8 storeys).
5. Commercial occupation (commercial building footprint) is expected to increase by approximately 20% between t_0 and t_f
6. Industrial occupation (industrial building footprint) is expected to increase by approximately 30% between t_0 and t_f
7. Non-residential buildings have the same probability distribution characteristics as residential buildings with respect to lrs, $code$ and h (height) attributes.

Appendix D. General assumptions for household and individual data generation in Tomorrowville

Household and individual population data for land-use zones with an avgIncome field value of "highIncome" are largely based on unpublished household survey information collected for Ward no 3. of the Lalitpur Metropolitan City (LMC-3; for which 2020/21 tax/revenue collection exceeded 50 million). Household and individual population data for land-use zones with an avgIncome field value of "midIncome" are largely based on unpublished household survey information collected for Ward no. 21 of the Lalitpur Metropolitan City (LMC-21; for which 2020/21 tax/revenue collection was between 10 and 50 million); any unavailable population information in this survey (such as the distribution of household number) was obtained from the Ward no 3. survey. Specific head of household details for land-use zones with avgIncome field values of either "highIncome" or "midIncome" were obtained from a national-level annual household survey [106], given a lack of available information at the ward level. Household and individual population data for land-use zones with an avgIncome field value of "lowIncomeA" or "lowIncomeB" are based on national-level

Table A.8
Methods and their details on social data disaggregation.

Method name: Iterative proportional fitting		
Method description: IPF is used to construct a sample that is consistent with known statistics of a target population. The method requires information about marginal distributions of individual attributes as well as a frequency cross-table of all attributes involved that defines their correlations. An iterative reweighting procedure is employed to fit the multi-dimensional cross-table conditional on the marginal distributions. The generated sample exactly matches the target marginal distributions of attributes and preserves the provided correlation structure.		
Advantages	Disadvantages	Key references
- It is simple and accurate - The efficiency of the algorithm enables large-scale problems to be solved with minimum computation overhead	- The method only yields fractions of individuals (rather than explicit individual agents) - In its original form, it cannot link individuals to household characteristics and cannot account for information that is missing in the correlation structure but exists in the real population	[109]
Method name: Iterative Proportional Updating (IPU)		
Method description: IPU is a hierarchical version of IPF, in which the multi-dimensional tables are fit using single joint weights that simultaneously account for individuals and households.		
Advantages	Disadvantages	Key references
- Simultaneously captures individual and household-level attributes, for more accurate characterisations of the overall synthetic population than IPF - Leverages more efficient data storage approaches than IPF	- The method can quickly become computationally burdensome - Can fail in extreme cases (i.e., when all persons of certain types fall into a single household type and can only capture nested relationships (i.e., person-residence associations but not person-workplace associations, for a given household)	[110–112]
Method name: Monte Carlo Simulation Methods (MCS)		
Method description: MCS covers a variety of simulation-based techniques, including Markov Chain Monte Carlo simulation. These approaches draw synthetic samples of individuals from conditional distributions of individual attributes, which form a partial view of the true population structure. The empirical distribution of the simulated population is as close as possible to the unique joint distributions in the actual population.		
Advantages	Disadvantages	Key references
- Generally produces more accurate results than IPF, across different scales and sampling rates - Only requires sample data; information on marginal distributions of attributes is not crucial - Sample from the distribution rather than simple cloning	- Not capable of synthesising a full household of individuals in their original form	[113]
Method name: Combinatorial Optimisation (CO)		
Method description: CO covers a variety of approaches, including simulated annealing. The procedure starts with a random subset of households/individuals from the provided micro sample. The selected households/individuals are iteratively replaced to improve their fit to the target marginal distributions; if a change improves the fit, then the swap is accepted. The performance of the fit is continuously assessed and the algorithm terminates when the most accurate synthetic population is obtained.		
Advantages	Disadvantages	Key references
- Generally produces more accurate results than IPF - Results in entire individuals or households (rather than fractions)	- Cannot guarantee the optimal solution is reached - Requires excessive computational time as population size grows	[114,115]
Method name: Hierarchical Models (HM)		
Method description: These methods can cover any of those mentioned in previous tables (i.e., IPF, CO, and MCS) and their adaptation to account for both individual and household attributes, associating these interdependent attributes in the most optimal way possible. HM also covers Bayesian Updating approaches, which use directed acyclical graphs to describe the conditional distribution of random variables, and respect the hierarchical structure of households and individuals. Furthermore, HM covers Hierarchical Mixture Modelling. Little attention has been paid to reproduce cross-level and within-household associations.		
Advantages	Disadvantages	Key references
- Ensure that the generated individuals/households respect the structure of each other - See previous sections for additional advantages associated with the underlying methods	- Bayesian frameworks are challenging from a computational perspective (it is better for the number of considered attributes to be small) - See previous sections for additional disadvantages associated with the underlying methods	[116,117]

(*continued on next page*)

Table A.8 (*continued*).

Method name: Deep Generative Modelling (DGM)		
Method description: DGM covers a variety of methods, including Variable Auto Encoders (an unsupervised model based on a deep artificial neural network) and General Adversarial Networks. Each person is represented as a vector of random variables, and the objective is to learn the joint distribution of all individuals in the sample.		
Advantages	Disadvantages	Key references
- Can be used to generate a consistent synthetic population of households and individuals with many attributes - Both categorical and numerical variables can be modelled jointly	- Requires advanced computational expertise - Cannot make direct use of marginal population count data - May lead to overfitting when trained with small datasets	[118–121]

Table B.9
Methods and their details on activity modelling.

Paper name	Activity model category	Locations captured	Methods used	Brief description	Information required
[133]	(1)	Home, Workplace, and Discretionary Activities (in particular)	Space–time prism constraint and its 2D projection (i.e., potential path area)	Uses space–time constraints to derive a potential path area (PPA) containing a maximum number of potential alternative discretionary activity locations. Assumes that choices for discretionary activity locations are spatially related to fixed locations of home and workplace.	Household and travel survey, providing information on trips as well as socio-demographic and economic characteristics of household members
[134]	(1)	Home, Workplace, School, and Discretionary Activities	Space–time prism constraint and its 2D projection (i.e., PPA)	Develops a methodology for applying the concept of the space–time prism to PPA development and appropriate location choice formation, considering the dynamic nature of the urban environment	Household travel survey, land parcel database, employment database (number of persons employed in each industry), database of business establishments
[135]	(1) & (2)	Home, Workplace, School, and Discretionary Activities	Random utility maximisation assumption	Proposes a framework for activity-travel scheduling of workers and students. The framework simultaneously accounts for space–time constraints (which determine the PPA of feasible locations for the next activity) and microeconomic theory of random utility maximising choice behaviour.	Household travel survey, including activity schedules and household/individual socio-economic variables
[136]	(2)	Workplace	Utility maximisation, using a multinomial logit model	Leverages a utility function to efficiently assign commuters to workplaces while respecting individual commuter preferences	Household travel survey information (including revealed preference data), spatial information on workplace locations
[137]	(3)	Home, Workplace, School, University	Gravity modelling	Examines commuting patterns by developing a gravity model that accounts for origin–destination fixed effects	Microdata files of workplace, school, and university locations
[138]	(1) & (3)	Discretionary Activities	Bayesian networks	Models the location choices of discretionary activities, by combining space–time constraints and the heuristics of individuals' location selection based on a Bayesian network hybrid learning algorithm	One-day travel diary, including socio-demographic characteristics, trip purpose, and zone destination

Table C.10
$p(occ|poly_t)$ distributions for Tomorrowville. Res: Residential; Com: Commercial; Ind: Industrial.

$poly_t$	TV0			TV50		
	Res	Com	Ind	Res	Com	Ind
Agriculture	1.00	0.00	0.00	1.00	0.00	0.00
CommercialResidential	0.00	1.00	0.00	0.00	1.00	0.00
HistoricalPreservation	0.70	0.30	0.00	0.70	0.30	0.00
Industry	0.00	0.00	1.00	0.00	0.00	1.00
CityCentre	0.00	1.00	0.00	0.00	1.00	0.00
Residential (lowIncomeA)	0.90	0.10	0.00	1.00	0.00	0.00
Residential (lowIncomeB)	0.90	0.10	0.00	1.00	0.00	0.00
Residential (midIncome)	0.60	0.40	0.00	1.00	0.00	0.00
Residential (highIncome)	0.80	0.20	0.00	1.00	0.00	0.00

Table D.11
Empirical statistics and probability distributions used for generating household- and individual-level attribute data.

Variable/Function	Value (avgIncome: highIncome)	Value (avgIncome: middleIncome)	Value (avgIncome: lowIncomeA)	Value (avgIncome: lowIncomeB)
$age_{school_{min}}$	2	2	6	6
$age_{school_{max}}$	17	17	17	17
age_{old}	50	50	50	50
$age_{hh,min}$	20	20	18	18
age_{adult}	18	18	18	18
$age_{university}$	25	25	25	25
$p(school\|age_{school_{min}} \le age \le age_{school_{max}}, gender)$	$p(school\|male) = 1$ $p(school\|female) = 1$	$p(school\|male) = 1$ $p(school\|female) = 1$	$p(school\|male) = 0.85$ $p(school\|female) = 0.87$ $p(school\|male, eduAttStat_{hh} > primary) = 1$ $p(school\|female, eduAttStat_{hh} > primary) = 1$	$p(school\|male) = 0.85$ $p(school\|female) = 0.87$ $p(school\|male, eduAttStat_{hh} > primary) = 1$ $p(school\|female, eduAttStat_{hh} > primary) = 1$
$p(nInd)$	$p(nInd = 1) = 0.224$ $p(nInd = 2) = 0.06$ $p(nInd = 3) = 0.113$ $p(nInd = 4) = 0.237$ $p(nInd = 5) = 0.149$ $p(nInd = 6) = 0.083$ $p(nInd = 7) = 0.032$ $p(nInd = 8) = 0.032$ $p(nInd = 9) = 0.07$	$p(nInd = 1) = 0.024$ $p(nInd = 2) = 0.053$ $p(nInd = 3) = 0.131$ $p(nInd = 4) = 0.258$ $p(nInd = 5) = 0.186$ $p(nInd = 6) = 0.130$ $p(nInd = 7) = 0.065$ $p(nInd = 8) = 0.064$ $p(nInd = 9) = 0.090$	$p(nInd = 1) = 0.222$ $p(nInd = 2) = 0.202$ $p(nInd = 3) = 0.172$ $p(nInd = 4) = 0.152$ $p(nInd = 5) = 0.11$ $p(nInd = 6) = 0.066$ $p(nInd = 7) = 0.043$ $p(nInd = 8) = 0.016$ $p(nInd = 9) = 0.017$	$p(nInd = 1) = 0.222$ $p(nInd = 2) = 0.202$ $p(nInd = 3) = 0.172$ $p(nInd = 4) = 0.152$ $p(nInd = 5) = 0.11$ $p(nInd = 6) = 0.066$ $p(nInd = 7) = 0.043$ $p(nInd = 8) = 0.016$ $p(nInd = 9) = 0.017$
$p(age\|male)$	$p(age = 0–1) = 0.016$ $p(age = 2–9) = 0.065$ $p(age = 10–17) = 0.105$ $p(age = 18–19) = 0.026$ $p(age = 20–25) = 0.103$ $p(age = 26–29) = 0.069$ $p(age = 30–39) = 0.202$ $p(age = 40–49) = 0.162$ $p(age = 50+) = 0.253$	$p(age = 0–1) = 0.016$ $p(age = 2–9) = 0.065$ $p(age = 10–17) = 0.105$ $p(age = 18–19) = 0.026$ $p(age = 20–25) = 0.103$ $p(age = 26–29) = 0.069$ $p(age = 30–39) = 0.202$ $p(age = 40–49) = 0.162$ $p(age = 50+) = 0.253$	$p(age = 0–5) = 0.15$ $p(age = 6–17) = 0.20$ $p(age = 18–25) = 0.19$ $p(age = 26–49) = 0.42$ $p(age = 50+) = 0.04$	$p(age = 0–5) = 0.15$ $p(age = 6–17) = 0.20$ $p(age = 18–25) = 0.19$ $p(age = 26–49) = 0.42$ $p(age = 50+) = 0.04$

(continued on next page)

household population and housing characteristics [139] as well as statistics for the Kibera slum in Nairobi [140]. Specific head of household details for these land-use zones were obtained from a country-level (Kenya) open dataset [141], given a lack of information at a more detailed resolution.

Employment and workplace location information for land-use zones with an avgIncome field value of "highIncome" were obtained from a national-level Nepalese labour force survey (due to a lack of availability of more granular information) [106]. Employment and workplace location information for land-use zones with an avgIncome field value of "midIncome" are largely based on the results of a relevant survey of LMC-21. Employment and workplace location information for land-use zones with an avgIncome field value of "lowIncomeA" or "lowIncomeB" were largely obtained from statistics for the Kibera slum in Nairobi [140] as well as Kenya-wide reporting data [141] and international economic data on slums [142].

Educational information for land-use zones with an avgIncome field value of "highIncome" or "midIncome" is based on a combination of expert judgement, data obtained from [143], and conversations with the ward chair of LMC-3. Educational

Table D.11 (*continued*).

Variable/Function	Value (avgIncome: highIncome)	Value (avgIncome: middleIncome)	Value (avgIncome: lowIncomeA)	Value (avgIncome: lowIncomeB)
$p(age\|female)$	$p(age = 0–1) = 0.014$ $p(age = 2–9) = 0.056$ $p(age = 10–17) = 0.088$ $p(age = 18–19) = 0.022$ $p(age = 20–25) = 0.102$ $p(age = 26–29) = 0.068$ $p(age = 30–39) = 0.19$ $p(age = 40–49) = 0.17$ $p(age = 50+) = 0.29$	$p(age = 0–1) = 0.014$ $p(age = 2–9) = 0.056$ $p(age = 10–17) = 0.088$ $p(age = 18–19) = 0.022$ $p(age = 20–25) = 0.102$ $p(age = 26–29) = 0.068$ $p(age = 30–39) = 0.19$ $p(age = 40–49) = 0.17$ $p(age = 50+) = 0.29$	$p(age = 0–5) = 0.175$ $p(age = 6–17) = 0.22$ $p(age = 18–25) = 0.26$ $p(age = 26–49) = 0.32$ $p(age = 50+) = 0.025$	$p(age = 0–5) = 0.175$ $p(age = 6–17) = 0.22$ $p(age = 18–25) = 0.26$ $p(age = 26–49) = 0.32$ $p(age = 50+) = 0.025$
$p(gender_{hh})$	$p(male) = 0.77$ $p(female) = 0.23$	$p(male) = 0.84$ $p(female) = 0.16$	$p(male) = 0.8$ $p(female) = 0.2$	$p(male) = 0.8$ $p(female) = 0.2$
$p(age_{hh}\|male)$	$p(age = 20–25) = 0.131$ $p(age = 26–29) = 0.087$ $p(age = 30–39) = 0.256$ $p(age = 40–49) = 0.205$ $p(age = 50+) = 0.321$	$p(age = 20–25) = 0.131$ $p(age = 26–29) = 0.087$ $p(age = 30–39) = 0.256$ $p(age = 40–49) = 0.205$ $p(age = 50+) = 0.321$	$p(age = 18–25) = 0.29$ $p(age = 26–49) = 0.65$ $p(age = 50+) = 0.06$	$p(age = 18–25) = 0.29$ $p(age = 26–49) = 0.65$ $p(age = 50+) = 0.06$
$p(age_{hh}\|female)$	$p(age = 20–25) = 0.124$ $p(age = 26–29) = 0.083$ $p(age = 30–39) = 0.232$ $p(age = 40–49) = 0.207$ $p(age = 50+) = 0.354$	$p(age = 20–25) = 0.124$ $p(age = 26–29) = 0.083$ $p(age = 30–39) = 0.232$ $p(age = 40–49) = 0.207$ $p(age = 50+) = 0.354$	$p(age = 18–25) = 0.43$ $p(age = 26–49) = 0.53$ $p(age = 50+) = 0.04$	$p(age = 18–25) = 0.43$ $p(age = 26–49) = 0.53$ $p(age = 50+) = 0.04$
$p(eduAttStat\|age > age_{adult}, male)$	$p(eduAttStat \le primary) = 0.33$ $p(eduAttStat = secondary) = 0.22$ $p(eduAttStat = university) = 0.45$	$p(eduAttStat \le primary) = 0.38$ $p(eduAttStat = secondary) = 0.40$ $p(eduAttStat = university) = 0.22$	$p(eduAttStat \le primary) = 0.7$ $p(eduAttStat = secondary) = 0.27$ $p(eduAttStat = university) = 0.03$	$p(eduAttStat \le primary) = 0.70$ $p(eduAttStat = secondary) = 0.27$ $p(eduAttStat = university) = 0.03$
$p(eduAttStat\|age \ge age_{adult}, female)$	$p(eduAttStat \le primary) = 0.43$ $p(eduAttStat = secondary) = 0.24$ $p(eduAttStat = university) = 0.33$	$p(eduAttStat \le primary) = 0.57$ $p(eduAttStat = secondary) = 0.23$ $p(eduAttStat = university) = 0.19$	$p(eduAttStat \le primary) = 0.79$ $p(eduAttStat = secondary) = 0.18$ $p(eduAttStat = university) = 0.03$	$p(eduAttStat \le primary) = 0.79$ $p(eduAttStat = secondary) = 0.18$ $p(eduAttStat = university) = 0.03$
$p(wp\|male)$	$p(professional) = 0.45$ $p(professional\|eduAttStat = university) = 1$ $p(industry) = 0.32$ $p(agriculture) = 0.13$ $p(unemployed) = 0.10$	$p(professional) = 0.78$ $p(professional\|eduAttStat = university) = 1$ $p(industry) = 0.04$ $p(agriculture) = 0.09$ $p(middleInc_{res}) = 0.01$ $p(unemployed) = 0.08$	$p(professional) = 0.35$ $p(professional\|eduAttStat = university) = 1$ $p(lowIncA) = 0.38$ $p(unemployed) = 0.27$	$p(professional) = 0.35$ $p(lowIncB\|eduAttStat = university) = 1$ $p(lowIncB) = 0.38$ $p(unemployed) = 0.27$
$p(wp\|female)$	$p(professional) = 0.42$ $p(professional\|eduAttStat = university) = 1$ $p(industry) = 0.16$ $p(agriculture) = 0.29$ $p(unemployed) = 0.13$	$p(professional) = 0.53$ $p(professional\|eduAttStat = university) = 1$ $p(agriculture) = 0.02$ $p(middleInc_{res}) = 0.34$ $p(unemployed) = 0.11$	$p(professional) = 0.15$ $p(professional\|eduAttStat = university) = 1$ $p(lowIncA) = 0.35$ $p(unemployed) = 0.5$	$p(professional) = 0.05$ $p(lowIncB\|eduAttStat = university) = 1$ $p(lowIncB) = 0.4$ $p(unemployed) = 0.55$

information for land-use zones with an avgIncome field value of "lowIncomeA" or "lowIncomeB" was obtained from reports and household-survey-level information for the Kibera slum in Nairobi [105] as well as Kenya-wide reporting data [144]. All other information (including empirical information related to correlations between different attributes) is based on expert judgement of experienced social scientists. Empirical statistics and probability distributions used for generating household- and individual-level attribute data are provided in Table D.11. Further assumptions are:

1. Where required, polygon- and gender-specific ages are assumed to be evenly distributed within a given provided age range (e.g., if someone is assigned an age range of 10–19, they have a 10% chance of being any single discrete age between 10 and 19, inclusive).
2. It is assumed that households with $nInd \ge 9$ have exactly nine members.
3. It is assumed that household and individual data do not change over time.

Note that Individuals who work in "professional" (i.e., professional or service-type) workplace locations are equally likely to be assigned a workplace building across: "CityCentre", "CommercialResidential", "Residential (highIncome)" (commercial buildings only), and "Residential (midIncome)" (commercial buildings only). $lowIncA$ denotes "Residential (lowIncomeA)", $lowIncB$ denotes "Residential (lowIncomeB)", $unemployed$ denotes unemployment, and $middleInc_{res}$ represents "Residential (midIncome)" (residential buildings only). Note that conditionals are removed from some of the probability expressions in the table, for brevity. Note that probability distributions $p(A)$ are related to conditional versions of $p(A|x)$, as follows:

$$p(A|x)\,r_x + p(A|\bar{x})\,r_{\bar{x}} = p(A) \tag{D.1}$$

where r_x is the proportion of the population of interest associated with the property x, $r_{\bar{x}}$ is the remaining proportion of that population (i.e., those not associated with the property x). For example, if A denotes adults who work in professional workplace locations, and x implies adults with a university (or college-level degree), then $\bar{x}$ implies adults with $eduAttStat \leq secondary$.

References

[1] U. Nations, D. of Economic, S. Affairs, P. Division, World population prospects 2019 highlights, 2019.
[2] J. Twigg, Disaster Risk Reduction, Overseas Development Institute, London, 2015.
[3] T. Elmqvist, C.L. Redman, S. Barthel, R. Costanza, History of urbanization and the missing ecology, in: Urbanization, Biodiversity and Ecosystem Services: Challenges and Opportunities: A Global Assessment, 2013, pp. 13–30.
[4] M. Moore, P. Gould, B.S. Keary, Global urbanization and impact on health, Int. J. Hygiene Environ. Health 206 (2003) 269–278, http://dx.doi.org/10.1078/1438-4639-00223.
[5] U. Nations, Sendai framework for disaster risk reduction, (ISSN: 08203946) 2015.
[6] U. Nations, Transforming our world: The 2030 agenda for sustainable development, 2015.
[7] IPCC, Climate change 2021 the physical science basis, (ISSN: 1003-3572) 2021, http://dx.doi.org/10.3724/sp.j.7103161536.
[8] W.M. Organization, WMO atlas of mortality and economic losses from weather, climate and water extremes (1970–2019), 2021.
[9] C. Mesta, G. Cremen, C. Galasso, Urban growth modelling and social vulnerability assessment for a hazardous kathmandu valley, Sci. Rep. (2022) http://dx.doi.org/10.1038/s41598-022-09347-x.
[10] E.L.F. Schipper, F. Thomalla, G. Vulturius, M. Davis, K. Johnson, Linking disaster risk reduction, climate change and development, Int. J. Disaster Resil. Built Environ. (2016).
[11] UNDRR, Global assessment report, ISBN: 9789210041805, 2019, p. 472.
[12] IPCC, Special report of the intergovernmental panel on climate change, managing the risks of extreme events and disasters to advance climate change adaptation, 2012.
[13] G.M. Raymond, The role of the physical urban planner, 1978, pp. 1–13.
[14] S.N. Suri, C. Johnson, B. Lipietz, Words into action guidelines: Implementation guide for land use and urban planning, 2020.
[15] M. Huxley, A. Inch, Urban planning, Int. Encyclopedia Hum. Geogr. (2020) 87–92, http://dx.doi.org/10.1016/B978-0-08-102295-5.10228-8.
[16] C. Galasso, J. McCloskey, M. Pelling, M. Hope, C.J. Bean, G. Cremen, R. Guragain, U. Hancilar, J. Menoscal, K. Mwang'a, J. Phillips, D. Rush, H. Sinclair, Editorial. Risk-based, pro-poor urban design and planning for Tomorrow's cities, Int. J. Disaster Risk Reduct. 58 (2021) 102158, http://dx.doi.org/10.1016/j.ijdrr.2021.102158.
[17] G. Cremen, C. Galasso, J. McCloskey, A simulation-based framework for earthquake risk-informed and people-centered decision making on future urban planning, Earth's Future 10 (2022) http://dx.doi.org/10.1029/2021EF002388.
[18] G. Cremen, C. Galasso, J. McCloskey, A. Barcena, M. Creed, M.E. Filippi, R. Gentile, L. Jenkins, M. Kalaycioglu, E.Y. Mentese, et al., A state-of-the-art decision-support environment for risk-sensitive and pro-poor urban planning and design in Tomorrow's cities, Int. J. Disaster Risk Reduct. (2022) 103400.
[19] M.P. Heris, A. Middel, B. Muller, Impacts of form and design policies on urban microclimate: Assessment of zoning and design guideline choices in urban redevelopment projects, Landsc. Urban Plan. 202 (2020) 103870.
[20] M. Banba, R. Shaw, Land Use Management in Disaster Risk Reduction, Springer, 2017.
[21] A.M. Hersperger, E. Oliveira, S. Pagliarin, G. Palka, P. Verburg, J. Bolliger, S. Grădinaru, Urban land-use change: The role of strategic spatial planning, Global Environ. Change 51 (2018) 32–42.
[22] D. Farr, Sustainable Urbanism: Urban Design with Nature, John Wiley & Sons, 2011.
[23] U. Avin, R. Goodspeed, Using exploratory scenarios in planning practice: A spectrum of approaches, J. Am. Plan. Assoc. 86 (4) (2020) 403–416.
[24] A. Chakraborty, A. McMillan, Scenario planning for urban planners: Toward a practitioner's guide, J. Am. Plan. Assoc. 81 (1) (2015) 18–29.
[25] L.D. Hopkins, M. Zapata, Engaging the Future: Forecasts, Scenarios, Plans, and Projects, Lincoln Institute of Land Policy, 2007.
[26] G. Abou Jaoude, O. Mumm, V.M. Carlow, An overview of scenario approaches: A guide for urban design and planning, J. Plan. Lit. (2022) 08854122221083546.
[27] L. Börjeson, M. Höjer, K.-H. Dreborg, T. Ekvall, G. Finnveden, Scenario types and techniques: towards a user's guide, Futures 38 (7) (2006) 723–739.
[28] J. Stapleton, Exploratory scenario planning, how to navigate an uncertain future, 2006.
[29] E.J. Roberts, Exploratory Scenario Planning: Lessons Learned in the Field, Lincoln Institute of Land Policy., 2014.
[30] J. Star, E.L. Rowland, M.E. Black, C.A. Enquist, G. Garfin, C.H. Hoffman, H. Hartmann, K.L. Jacobs, R.H. Moss, A.M. Waple, Supporting adaptation decisions through scenario planning: Enabling the effective use of multiple methods, Clim. Risk Manage. 13 (2016) 88–94.
[31] B. Rogers, N. Bertram, B. Gersonius, A. Gunn, R. Löwe, C. Murphy, R. Pasman, M. Radhakrishnan, C. Urich, T. Wong, et al., An interdisciplinary and catchment approach to enhancing urban flood resilience: a Melbourne case, Phil. Trans. R. Soc. A 378 (2168) (2020) 20190201.
[32] T. McPhearson, D. Haase, N. Kabisch, Å. Gren, Advancing understanding of the complex nature of urban systems, 2016.
[33] R.E. Klosterman, Lessons learned about planning: Forecasting, participation, and technology, J. Am. Plan. Assoc. 79 (2) (2013) 161–169.
[34] A. Karvonen, V. Cvetkovic, P. Herman, K. Johansson, H. Kjellström, M. Molinari, M. Skoglund, The 'new urban science': towards the interdisciplinary and transdisciplinary pursuit of sustainable transformations, Urban Transform. 3 (1) (2021) 1–13.
[35] P. Rode, A. Gomes, M. Adeel, F. Sajjad, A. Koch, S.M. Murshed, Between abundance and constraints: the natural resource equation of Asia's diverging, higher-income city models, Land 9 (11) (2020) 426.
[36] G.A. Riddell, H. van Delden, H.R. Maier, A.C. Zecchin, Exploratory scenario analysis for disaster risk reduction: Considering alternative pathways in disaster risk assessment, Int. J. Disaster Risk Reduct. 39 (2019) 101230.
[37] X. Wu, Z. Wang, S. Guo, W. Liao, Z. Zeng, X. Chen, Scenario-based projections of future urban inundation within a coupled hydrodynamic model framework: a case study in Dongguan city, China, J. Hydrol. 547 (2017) 428–442.
[38] M. Mokrech, R. Nicholls, J. Richards, C. Henriques, I. Holman, S. Shackley, Regional impact assessment of flooding under future climate and socio-economic scenarios for east Anglia and north west England, Clim. Change 90 (1) (2008) 31–55.

[39] U. Gigler, H. Hemis, Innovative approaches to integrative energy planning–experiences and results from the EU project urban learning, in: REAL CORP 2018–EXPANDING CITIES–DIMINISHING SPACE. are ''Smart Cities'' the Solution Or Part of the Problem of Continuous Urbanisation Around the Globe? Proceedings of 23rd International Conference on Urban Planning, Regional Development and Information, CORP–Compentence Center of Urban and Regional Planning, 2018, pp. 613–622.
[40] T. Yigitcanlar, S. Teriman, Rethinking sustainable urban development: towards an integrated planning and development process, Int. J. Environ. Sci. Technol. (2015) http://dx.doi.org/10.1007/s13762-013-0491-x.
[41] P. Berke, E.J. Kaiser, Urban Land Use Planning, University of Il linois Press, 2006.
[42] G. Zuccaro, M. Leone, C. Martucci, Future research and innovation priorities in the field of natural hazards, disaster risk reduction, disaster risk management and climate change adaptation: A shared vision from the ESPREssO project, Int. J. Disaster Risk Reduct. 51 (2020) 101783.
[43] B.M. Sleeter, N.J. Wood, C.E. Soulard, T.S. Wilson, Projecting community changes in hazard exposure to support long-term risk reduction: A case study of tsunami hazards in the US Pacific northwest, Int. J. Disaster Risk Reduct. 22 (2017) 10–22.
[44] L. Peng, J. Xia, Z. Li, C. Fang, X. Deng, Spatio-temporal dynamics of water-related disaster risk in the Yangtze river economic belt from 2000 to 2015, Resour. Conserv. Recy. 161 (2020) 104851.
[45] G. Cremen, C. Galasso, J. McCloskey, Modelling and quantifying Tomorrow's risks from natural hazards, Sci. Total Environ. 817, 152552, http://dx.doi.org/10.1016/j.scitotenv2021.152552.
[46] F.E.M. Agency, Hazus inventory technical manual, hazus 4.2 service pack 3, 2021.
[47] V. Silva, D. Amo-Oduro, A. Calderon, C. Costa, J. Dabbeek, V. Despotaki, L. Martins, M. Pagani, A. Rao, M. Simionato, et al., Development of a global seismic risk model, Earthq. Spectra 36 (1_suppl) (2020) 372–394.
[48] 2021. [link]. URL https://github.com/OasisLMF/OpenDataStandards.
[49] 2021. [link]. URL https://docs.riskdatalibrary.org/exposure.html.
[50] A. Calderón, V. Silva, Exposure forecasting for seismic risk estimation: Application to Costa Rica, Earthq. Spectra 37 (3) (2021) 1806–1826.
[51] C. Linard, A.J. Tatem, M. Gilbert, Modelling spatial patterns of urban growth in Africa, Appl. Geogr. 44 (2013) 23–32.
[52] S. Balica, Q. Dinh, I. Popescu, Vulnerability and exposure in developed and developing countries: Large-scale assessments, in: Hydro-Meteorological Hazards, Risks and Disasters, Elsevier, 2015, pp. 125–162.
[53] L. Yang, J. Scheffran, H. Qin, Q. You, Climate-related flood risks and urban responses in the pearl River Delta, China, Reg. Environ. Change 15 (2) (2015) 379–391.
[54] A. Fekete, A. Asadzadeh, M. Ghafory-Ashtiany, K. Amini-Hosseini, C. Hetkämper, M. Moghadas, A. Ostadtaghizadeh, A. Rohr, T. Kötter, Pathways for advancing integrative disaster risk and resilience management in Iran: Needs, challenges and opportunities, Int. J. Disaster Risk Reduct 49 (2020) 101635.
[55] A.H. Thieken, S. Kienzler, H. Kreibich, C. Kuhlicke, M. Kunz, B. Mühr, M. Müller, A. Otto, T. Petrow, S. Pisi, et al., Review of the flood risk management system in Germany after the major flood in 2013, Ecol. Soc. 21 (2) (2016).
[56] D. Piper, M. Kunz, F. Ehmele, S. Mohr, B. Mühr, A. Kron, J. Daniell, Exceptional sequence of severe thunderstorms and related flash floods in may and june 2016 in Germany–part 1: Meteorological background, Nat. Hazards Earth Syst. Sci. 16 (12) (2016) 2835–2850.
[57] A. Fekete, S. Sandholz, Here comes the flood, but not failure? Lessons to learn after the heavy rain and pluvial floods in Germany 2021, Water 13 (21) (2021) 3016.
[58] E.L. Bohensky, A.M. Leitch, Framing the flood: a media analysis of themes of resilience in the 2011 brisbane flood, Reg. Environ. Change 14 (2) (2014) 475–488.
[59] C. Johnson, Urban and regional planning and disaster, in: The Routledge Handbook of Hazards and Disaster Risk Reduction, Routledge, 2012, pp. 641–651.
[60] A. March, J. Leon, Urban planning for disaster risk reduction: Establishing 2nd wave criteria, in: K. Ruming, B. Randolph, N. Gurran (Eds.), State of Australian Cities, State of Australian Cities Research Network, Sydney, NSW, 2013.
[61] S. Wallin, et al., Managing urban complexity-participatory planning, self-organization and co-production of urban space, 2019.
[62] J. Friedmann, Planning in the Public Domain: From Knowledge to Action, Princeton University Press, 1987.
[63] J. León, A. March, An urban form response to disaster vulnerability: Improving tsunami evacuation in Iquique, Chile, Environ. Plan. B: Plann. Des. 43 (5) (2016) 826–847.
[64] D. King, Y. Gurtner, A. Firdaus, S. Harwood, A. Cottrell, Land use planning for disaster risk reduction and climate change adaptation: Operationalizing policy and legislation at local levels, Int. J. Disaster Resil. Built Environ. (2016).
[65] M.E. Filippi, A. Barcena, R.Š. Trogrlić, G. Cremen, E. Menteşe, R. Gentile, M. Creed, L. Jenkins, M. Kalaycioglu, D. Poudel, et al., Interdisciplinarity in practice: Reflections from early-career researchers developing a risk-informed decision support environment for Tomorrow's cities, Int. J. Disaster Risk Reduct. 85 (2023) 103481.
[66] V. Gandhi, Vector data, in: S. Shekhar, H. Xiong, X. Zhou (Eds.), Encyclopedia of GIS, Springer International Publishing, 2017, pp. 2411–2416, http://dx.doi.org/10.1007/978-3-319-17885-1_1438.
[67] Vector data model, 2013, URL http://wiki.gis.com/wiki/index.php/Vector_data_model.
[68] M.C. Lens, Zoning, land use, and the reproduction of urban inequality, Annu. Rev. Sociol. 48 (2022).
[69] D. Farr, Sustainable Urbanism: Urban Design with Nature, John Wiley & Sons, 2008.
[70] OECD, The governance of land use, policy highlights, 2017, URL https://www.oecd.org/cfe/regionaldevelopment/governance-of-land-use-policy-highlights.pdf.
[71] T.F.o.S.S. American Psychological Association, Report of the APA task force on socioeconomic status, 2007, URL https://www.apa.org/pi/ses/resources/publications/task-force-2006.pdf.
[72] E. Moro, D. Calacci, X. Dong, A. Pentland, Mobility patterns are associated with experienced income segregation in large US cities, Nature Commun. 12 (1) (2021) 1–10.
[73] P. Veneri, A. Comandon, M.-À. Garcia-López, M.N. Daams, What do divided cities have in common? An international comparison of income segregation, J. Reg. Sci. 61 (1) (2020) 162–188, http://dx.doi.org/10.1111/jors.12506.
[74] E. Berkes, R. Gaetani, Income Segregation and Rise of the Knowledge Economy, Rotman School of Management Working Paper (3423136), 2019.
[75] OECD, OECD regions and cities at a glance, 2018, URL https://doi.org/10.1787/reg_cit_glance-2018-en.
[76] V. Silva, S. Brzev, C. Scawthorn, C. Yepes, J. Dabbeek, H. Crowley, A building classification system for multi-hazard risk assessment, Int. J. Disaster Risk Sci. 13 (2) (2022) 161–177.
[77] [link]. URL https://wiki.openstreetmap.org/wiki/Main_Page.
[78] R. Gentile, G. Cremen, C. Galasso, L.T. Jenkins, V. Manandhar, E.Y. Menteşe, R. Guragain, J. McCloskey, Scoring, selecting, and developing physical impact models for multi-hazard risk assessment, Int. J. Disaster Risk Reduct. 82 (2022) 103365.
[79] [link]. URL https://wiki.openstreetmap.org/wiki/GED4ALL.
[80] UNDRR, Global assessment report on disaster risk reduction 2022: Our world at risk: Transforming governance for a resilient future, 2022, URL https://www.oecd.org/cfe/regionaldevelopment/governance-of-land-use-policy-highlights.pdf.
[81] A. Shepherd, T. Mitchell, K. Lewis, A. Lenhardt, L. Jones, L. Scott, R. Muir-Wood, The Geography of Poverty, Disasters and Climate Extremes in 2030, ODI London, 2013.
[82] S.L. Cutter, C. Finch, Temporal and spatial changes in social vulnerability to natural hazards, Proc. Natl. Acad. Sci. 105 (7) (2008) 2301–2306.

[83] J. Birkmann, O.D. Cardona, M.L. Carreño, A.H. Barbat, M. Pelling, S. Schneiderbauer, S. Kienberger, M. Keiler, D. Alexander, P. Zeil, et al., Framing vulnerability, risk and societal responses: the MOVE framework, Nat. Hazards 67 (2) (2013) 193–211.
[84] J. Ribot, Cause and response: vulnerability and climate in the anthropocene, J. Peasant Stud. 41 (5) (2014) 667–705.
[85] F.I. Rivera, N. Kapucu, Disaster Vulnerability, Hazards and Resilience, Springer International Publishing, Switzerland, 2015.
[86] D.G. Aliaga, P.A. Rosen, D.R. Bekins, Style grammars for interactive visualization of architecture, IEEE Trans. Vis. Comput. Graphics 13 (4) (2007) 786–797.
[87] D.S. Ebert, F.K. Musgrave, D. Peachey, K. Perlin, S. Worley, Texturing & Modeling: a Procedural Approach, Morgan Kaufmann, 2003.
[88] B. Watson, P. Müller, O. Veryovka, A. Fuller, P. Wonka, C. Sexton, Procedural urban modeling in practice, IEEE Comput. Graph. Appl. 28 (3) (2008) 18–26.
[89] G. Alomía, D. Loaiza, C. Zúñiga, X. Luo, R. Asorey-Cacheda, Procedural modeling applied to the 3D city model of bogota: a case study, Virtual Real. Intell. Hardw. 3 (5) (2021) 423–433.
[90] S.A. Biancardo, A. Capano, S.G. de Oliveira, A. Tibaut, Integration of BIM and procedural modeling tools for road design, Infrastructures 5 (4) (2020) 37.
[91] S.A. Groenewegen, R.M. Smelik, K.J. de Kraker, R. Bidarra, Procedural city layout generation based on urban land use models, in: Short Paper Proceedings of Eurographics 2009, 2009.
[92] G. Kelly, H. McCabe, A survey of procedural techniques for city generation, ITB J. 14 (3) (2006) 342–351.
[93] Y.I. Parish, P. Müller, Procedural modeling of cities, in: Proceedings of the 28th annual conference on Computer graphics and interactive techniques, 2001, pp. 301–308.
[94] T. Lechner, B. Watson, P. Ren, U. Wilensky, S. Tisue, M. Felsen, Procedural modeling of land use in cities, 2004.
[95] T. Lechner, P. Ren, B. Watson, C. Brozefski, U. Wilenski, Procedural modeling of urban land use, in: ACM SIGGRAPH 2006 Research Posters, 2006, pp. 135–es.
[96] R.M. Smelik, T. Tutenel, R. Bidarra, B. Benes, A survey on procedural modelling for virtual worlds, in: Computer Graphics Forum, Vol. 33, Wiley Online Library, 2014, pp. 31–50.
[97] J. Freiknecht, W. Effelsberg, A survey on the procedural generation of virtual worlds, Multimodal Technol. Interact. 1 (4) (2017) 27.
[98] J. Maantay, A. Maroko, Assessing population at risk: Areal interpolation and dasymetric mapping, in: The Routledge Handbook of Environmental Justice, Routledge, 2017, pp. 190–206.
[99] J. Baynes, A. Neale, T. Hultgren, Improving intelligent dasymetric mapping population density estimates at 30-meter resolution for the conterminous United States by excluding uninhabited areas, Earth Syst. Sci. Data Discuss. (2022) 1–24.
[100] J. Dabbeek, H. Crowley, V. Silva, G. Weatherill, N. Paul, C.I. Nievas, Impact of exposure spatial resolution on seismic loss estimates in regional portfolios, Bull. Earthq. Eng. 19 (14) (2021) 5819–5841.
[101] B. Mwau, A. Sverdlik, J. Makau, Urban Transformation and the Politics of Shelter, Tech. rep., IIED Working Paper, IIED, London, 2020.
[102] The Project for Assessment of Earthquake Disaster Risk for the Kathmandu Valley in Nepal-Final Report, Tech. rep., JICA, 2018.
[103] Project for Investigation of Damage Situation of Cultural Heritage in Nepal, Tech. rep., NRICP, 2016, 2016.
[104] Vision 2035 and Beyond 20 Years Strategic Development Master Plan (2015 - 2035) for Kathmandu Valley, Tech. rep., Kathmandu Valley Development Authority, 2016.
[105] Building Urban Flood Resilience Integrating Community Perspectives, Tech. rep., Konkuey Design Initiative, 2016.
[106] National Population and Housing Census 2012, Tech. rep., Government of Nepal, Central Bureau of Statistics, 2012.
[107] L.T. Jenkins, M.J. Creed, K. Tarbali, M. Muthusamy, R.Š. Trogrlić, J.C. Phillips, C.S. Watson, H.D. Sinclair, C. Galasso, J. McCloskey, Physics-based simulations of multiple natural hazards for risk-sensitive planning and decision making in expanding urban regions, Int. J. Disaster Risk Reduct. 84 (2023) 103338.
[108] E. Gencer, How to make cities more resilient a handbook for local government leaders, 2017.
[109] R.J. Beckman, K.A. Baggerly, M.D. McKay, Creating synthetic baseline populations, Transp. Res. A 30 (6) (1996) 415–429.
[110] M.-R. Namazi-Rad, R. Tanton, D. Steel, P. Mokhtarian, S. Das, An unconstrained statistical matching algorithm for combining individual and household level geo-specific census and survey data, Comput. Environ. Urban Syst. 63 (2017) 3–14.
[111] X. Ye, K. Konduri, R.M. Pendyala, B. Sana, P. Waddell, A methodology to match distributions of both household and person attributes in the generation of synthetic populations, in: 88th Annual Meeting of the Transportation Research Board, Washington, DC, 2009.
[112] D.R. Pritchard, E.J. Miller, Advances in Agent Population Synthesis and Application in an Integrated Land Use and Transportation Model, Tech. rep., 2009.
[113] B. Farooq, M. Bierlaire, R. Hurtubia, G. Flötteröd, Simulation based population synthesis, Transp. Res. B 58 (2013) 243–263.
[114] I. Abraham, M. Babaioff, S. Dughmi, T. Roughgarden, Combinatorial auctions with restricted complements, in: Proceedings of the 13th ACM Conference on Electronic Commerce, 2012, pp. 3–16.
[115] J. Ryan, H. Maoh, P. Kanaroglou, Population synthesis: Comparing the major techniques using a small, complete population of firms, Geogr. Anal. 41 (2) (2009) 181–203.
[116] D. Casati, K. Müller, P.J. Fourie, A. Erath, K.W. Axhausen, Synthetic population generation by combining a hierarchical, simulation-based approach with reweighting by generalized raking, Transp. Res. Rec. 2493 (1) (2015) 107–116.
[117] K. Müller, K.W. Axhausen, Hierarchical IPF: Generating a synthetic population for Switzerland, Arb. Verk. Raumplan. 718 (2011).
[118] Z. Aemmer, D. MacKenzie, Generative population synthesis for joint household and individual characteristics, Comput. Environ. Urban Syst. 96 (2022) 101852.
[119] B.F. Yameogo, P.O. Vandanjon, P. Gastineau, P. Hankach, Generating a two-layered synthetic population for french municipalities: Results and evaluation of four synthetic reconstruction methods, JASSS-J. Artif. Soc. Soc. Simul. 24 (2021) 27p.
[120] S.S. Borysov, J. Rich, F.C. Pereira, How to generate micro-agents? A deep generative modeling approach to population synthesis, Transp. Res. C 106 (2019) 73–97.
[121] M. Johnsen, O. Brandt, S. Garrido, F. Pereira, Population synthesis for urban resident modeling using deep generative models, Neural Comput. Appl. 34 (6) (2022) 4677–4692.
[122] S. Rasouli, H. Timmermans, Activity-based models of travel demand: promises, progress and prospects, Int. J. Urban Sci. 18 (1) (2014) 31–60, http://dx.doi.org/10.1080/12265934.2013.835118.
[123] H. Ballis, L. Dimitriou, Revealing personal activities schedules from synthesizing multi-period origin-destination matrices, Transp. Res. B 139 (2020) 224–258.
[124] J.W. Joubert, A. De Waal, Activity-based travel demand generation using Bayesian networks, Transp. Res. C 120 (2020) 102804.
[125] M.H. Hafezi, L. Liu, H. Millward, A time-use activity-pattern recognition model for activity-based travel demand modeling, Transportation 46 (4) (2019) 1369–1394.
[126] I. Saadi, A. Mustafa, J. Teller, M. Cools, Forecasting travel behavior using Markov chains-based approaches, Transp. Res. C 69 (2016) 402–417.
[127] C. Anda, S.A.O. Medina, K.W. Axhausen, Synthesising digital twin travellers: Individual travel demand from aggregated mobile phone data, Transp. Res. C 128 (2021) 103118.
[128] P. Liu, F. Liao, H.-J. Huang, H. Timmermans, Dynamic activity-travel assignment in multi-state supernetworks, Transp. Res. Procedia 7 (2015) 24–43.

[129] W. Yin, P. Murray-Tuite, S.V. Ukkusuri, H. Gladwin, An agent-based modeling system for travel demand simulation for hurricane evacuation, Transp. Res. C 42 (2014) 44–59.
[130] S. Jiang, J. Ferreira, M.C. Gonzalez, Activity-based human mobility patterns inferred from mobile phone data: A case study of Singapore, IEEE Trans. Big Data 3 (2) (2017) 208–219.
[131] S. Hörl, M. Balac, Synthetic population and travel demand for Paris and Île-de-France based on open and publicly available data, Transp. Res. C 130 (2021) 103291, http://dx.doi.org/10.1016/j.trc.2021.103291.
[132] N. Shoval, M. Isaacson, Sequence alignment as a method for human activity analysis in space and time, Ann. Assoc. Am. Geogr. 97 (2) (2007) 282–297.
[133] A. Justen, F.J. Martínez, C.E. Cortés, The use of space–time constraints for the selection of discretionary activity locations, J. Transp. Geogr. 33 (2013) 146–152.
[134] S.Y. Yoon, K. Deutsch, Y. Chen, K.G. Goulias, Feasibility of using time–space prism to represent available opportunities and choice sets for destination choice models in the context of dynamic urban environments, Transportation 39 (4) (2012) 807–823.
[135] K. Nurul Habib, A comprehensive utility-based system of activity-travel scheduling options modelling (CUSTOM) for worker's daily activity scheduling processes, Transp. A: Transp. Sci. 14 (4) (2018) 292–315.
[136] B.J. Vitins, A. Erath, K.W. Axhausen, Integration of a capacity-constrained workplace choice model: Recent developments and applications with an agent-based simulation in Singapore, Transp. Res. Rec. 2564 (1) (2016) 1–13.
[137] A. Ahrens, S. Lyons, Do rising rents lead to longer commutes? A gravity model of commuting flows in Ireland, Urban Stud. 58 (2) (2021) 264–279.
[138] T.-Y. Ma, S. Klein, Bayesian networks for constrained location choice modeling using structural restrictions and model averaging, Eur. J. Transp. Infrastr. Res. 18 (1) (2018).
[139] Kenya Demographic and Health Survey, Tech. rep., Central Bureau of Statistics, 2012.
[140] A. Desgroppes, S. Taupin, Kibera: The biggest slum in Africa? Cah. d'Afr. l'Est East Afr. Rev. (44) (2011) 23–33.
[141] Socio economic data of Kenya. URL https://kenya.opendataforafrica.org/SEDK2015/socio-economic-data-of-kenya-2011?region=1000010-nairobi&indicator=1012920-unemployment-rate-15-64-years.
[142] [link]. URL https://www.rockefellerfoundation.org/wp-content/uploads/Constrained-Opportunities-in-Slum-Economies.pdf.
[143] [link]. URL https://download.ei-ie.org/Docs/WebDepot/ECE_A_global_scenario_EN.PDF.
[144] Basic education sector analysis report, 2012, URL https://openjicareport.jica.go.jp/pdf/12083176.pdf.